On the Public Communication Needed to Achieve SK Capacity in the Multiterminal Source Model*

Manuj Mukherjee† Navin Kashyap† Yogesh Sankarasubramaniam‡

***Abstract*—The focus of this paper is on the public communication required for generating a maximal-rate secret key (SK) within the multiterminal source model of Csiszár and Narayan. Building on the prior work of Tyagi for the two-terminal scenario, we derive a lower bound on the communication complexity, R_{SK}, defined to be the minimum rate of public communication needed to generate a maximal-rate SK. It is well known that the minimum rate of communication for omniscience, denoted by R_{CO}, is an upper bound on R_{SK}. For the class of pairwise independent network (PIN) models defined on uniform hypergraphs, we show that a certain "Type $\mathcal{S}$" condition, which is verifiable in polynomial time, guarantees that our lower bound on R_{SK} meets the R_{CO} upper bound. Thus, PIN models satisfying our condition are "R_{SK}-maximal", indicating that the upper bound $R_{\mathrm{SK}} \leq R_{\mathrm{CO}}$ holds with equality. This allows us to explicitly evaluate R_{SK} for such PIN models. We also give several examples of PIN models that satisfy our Type $\mathcal{S}$ condition. Finally, we prove that for an arbitrary multiterminal source model, a stricter version of our Type $\mathcal{S}$ condition implies that communication from *all* terminals ("omnivocality") is needed for establishing a SK of maximum rate. For three-terminal source models, the converse is also true: omnivocality is needed for generating a maximal-rate SK only if the strict Type $\mathcal{S}$ condition is satisfied. However, for source models with four or more terminals, counterexamples exist that show that the converse does not hold in general.**

***Index Terms*—Communication complexity, information-theoretic security, PIN model, secret key generation**

I. Introduction

Maurer [1] and Ahlswede and Csiszár [2] independently introduced the problem of generating a secret key (SK) for a pair of terminals observing distinct, albeit correlated, components of a discrete memoryless multi-component source. The SK is to be generated by communicating interactively over a noiseless public channel, and it is to be kept secure from all passive eavesdroppers having access to the public channel. The problem was subsequently extended to a multiterminal setting by Csiszár and Narayan [3]. The Csiszár-Narayan model is now commonly referred to as the *multiterminal source model*. The quantity of interest in these papers, and indeed in much of the literature that followed on this topic [4], [5], [6], [7], is the *secret key capacity*, i.e., the supremum of the rates of SK that can be generated within this model. In the two-terminal case, an exact characterization of the SK capacity can be found in the original works of Maurer [1] and Ahlswede and Csiszár [2]. Csiszár and Narayan [3] later gave an elegant single-letter expression for SK capacity in the general multiterminal source model.

In all the aforementioned studies, the noiseless public channel is viewed as an unlimited free resource, and no attempt is made to restrict the amount of communication sent through it. Indeed, Csiszár and Narayan [3, Section VI] left open the question of determining the minimum rate of interactive public communication needed to achieve SK capacity. Tyagi [8] addressed this question in the two-terminal case, and gave an exact, although difficult to compute, characterization of the minimum rate of communication. In this paper, we extend some of Tyagi's ideas to the multiterminal setting, and apply them to give an explicit answer to Csiszár and Narayan's open question in some interesting special cases, namely, certain instances of the so-called pairwise independent network (PIN) model [5], [6].

A. Our Contributions

The primary focus of our work is on the following question: *What is the minimum rate of interactive public communication required to achieve SK capacity in the multiterminal source model?* We shall refer to the minimum rate of public communication as the *communication complexity*[1] of achieving SK capacity, denoted by R_{SK}. Csiszár and Narayan's original proof of the achievability of their single-letter expression for SK capacity [3, Theorem 1] used a (non-interactive) communication protocol that enabled "omniscience" at all terminals, which means that the communication over the public channel allows each terminal to recover the observations of all the other terminals. It follows from their results that R_{SK} is always upper bounded by R_{CO}, the minimum rate of interactive public communication required to achieve omniscience at all terminals. Furthermore, R_{CO} is given by the solution to a linear program [3, Proposition 1] (see (1) in Section II), so it can be computed efficiently. On the other hand, it is also well known that omniscience is not necessary for maximal-rate SK generation — see the remark following Theorem 1

†M. Mukherjee and N. Kashyap are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore. Email: {manuj,nkashyap}@ece.iisc.ernet.in.

‡Email: yogesh@gatech.edu

*This work was supported in part by a Swarnajayanti Fellowship awarded by the Department of Science and Technology, India. Parts of this work were presented at the 2014 IEEE International Symposium on Information Theory (ISIT 2014), Honolulu, Hawaii, USA, and at ISIT 2015, Hong Kong, China.

[1]Our use of "communication complexity" differs from the use prevalent in the theoretical computer science literature where, following [9], it refers to the total amount of communication, in bits, required to perform some distributed computation.

in [3], and also the proof of Theorem 3.2 in [4]. Indeed, it is not difficult to find examples where $R_{SK} \ll R_{\text{CO}}$; our Example IV.1 is one such. Thus, the sources for which we have $R_{\text{SK}} = R_{\text{CO}}$ constitute the worst-case sources in terms of communication complexity; we call such sources *R_{SK}-maximal*. We give a sufficient condition for a PIN model defined on a uniform hypergraph to be R_{SK}-maximal, and show that PIN models satisfying this condition do exist. For these PIN models, it is easy to explicitly compute R_{CO}, which then gives us an exact expression for R_{SK}. This is the first (non-trivial) explicit evaluation of R_{SK} to be found in the literature, for a multiterminal source model with more than two terminals. Interestingly, for PIN models defined on ordinary graphs (i.e., each edge is incident with only two vertices), our sufficient condition is also necessary, which gives us an exact characterization, decidable in polynomial time, of ordinary graph PIN models that are R_{SK}-maximal.

The main tool in our analysis is a lower bound on R_{SK} obtained via a multiterminal extension of Tyagi's work [8]. Tyagi's characterization of R_{SK} for two terminals [8, Theorem 3] was in terms of the minimum rate of an *interactive common information*, a type of Wyner common information (see [10]). In order to appropriately generalize these ideas, we propose extensions of conditional mutual information and Wyner common information to the setting of more than two terminals. With these new multiterminal definitions in hand, we essentially follow the approach in [8] to derive a lower bound on R_{SK} in terms of the minimum rate of a multiterminal analogue of interactive common information. As in the case of Tyagi's result for two terminals, an exact evaluation of this bound appears to be a difficult task even for simple source models such as Markov chains. However, unlike the two-terminal result, we are unable to show that our lower bound to R_{SK} is tight in general. Luckily, we are able to evaluate this bound exactly for certain PIN models as mentioned above, and the bound turns out to be tight in these cases as it matches the R_{CO} upper bound.

A secondary line of investigation carried out in this paper concerns the question of whether some terminals need to communicate at all in order to achieve SK capacity. It is well known that, in order to generate a maximal-rate SK in the two-terminal model, it is sufficient for only one terminal to communicate [1], [2], [3]. All this terminal has to do is convey its local observations to the other terminal at the least possible rate of communication required to do so. Thus, in the two-terminal setup, it is *never necessary* for both terminals to communicate to generate a capacity-achieving SK. Even in the case of more than two terminals, there are examples wherein not all terminals need to communicate — again, see the remark following Theorem 1 in [3]. However, as we will show in this paper, there are plenty of other examples where all terminals *must* communicate in order to achieve SK capacity. We coin the term "omnivocality" to describe the state when all terminals communicate. The problem of interest to us then is the following: *Characterize the instances of the multiterminal source model in which omnivocality is necessary for maximal-rate SK generation*. In this paper, we report some partial progress towards such a characterization.

In [7], Gohari and Anantharam considered the scenario where a subset of terminals is required to remain silent, and yet all the terminals must agree upon an SK using only the communication from terminals that are allowed to talk. They derived a linear programming formulation for the maximum SK rate achievable in this scenario. Observe that omnivocality is necessary for achieving SK capacity in a source model iff any one terminal not being allowed to communicate strictly lowers the maximum achievable SK rate for that model. This establishes a correspondence between the omnivocality condition and the Gohari-Anantharam scenario involving silent terminals. We use this correspondence to identify a sufficient condition under which omnivocality is necessary for achieving SK capacity in a source model with at least three terminals. We further show that in the case of exactly three terminals, our sufficient condition is also necessary. Based on this evidence, we had conjectured in [11] that our condition was always necessary and sufficient. Unfortunately, a counterexample has been given by Chan et al. [12] that shows that the condition is not necessary for four or more terminals. This has also been independently observed by Zhang et al. in [13].

B. Related work

Besides the work of Tyagi [8] that we have already mentioned, a few other recent papers have considered the SK generation problem from a communication complexity angle. The line of work that is perhaps most directly related to ours is that of Courtade and co-authors [14], [15], [16], which considers the *coded cooperative data exchange (CCDE)* problem [17] with the goal of generating an SK. This is, in essence, a single-shot version of the SK generation problem defined on hypergraph PIN models. Here, by "single-shot", we mean that each terminal sees only one realization of the component of the source available to it, as opposed to the Csiszár-Narayan setup within which each terminal sees a sequence of i.i.d. realizations. The single-shot SK capacity, i.e., the maximum size (as opposed to maximum rate) of an SK that can be generated, was evaluated in [14, Theorem 6]. The capacity achieving protocol used is a one-shot version of the communication-for-omniscience protocol of [3]. The follow-up works [15] and [16] addressed the issue of determining the minimum amount (again, as opposed to rate) of communication required to generate an SK of a particular size. However, this is done under a additional linearity requirement on the communication, i.e., the communication is required to be a linear function of the source outputs. Theorem 11 of [16] then gives an explicit characterization of what could be rightfully called the *linear communication complexity* of generating an SK of a given size, in terms of the minimum number of hyperedges of an "inherently τ-connected subhypergraph". It was further shown in [16, Theorem 4] that there exist hypergraph PIN models for which non-linear communication protocols for achieving (single-shot) SK capacity require lower amounts of communication than the linear communication complexity. It should be emphasized that our results do not make any linearity assumptions on the public communication.

In [18], Liu et al. study public communication for SK generation in another variant of the multiterminal source model.

The authors consider $m+1$ terminals observing correlated i.i.d. sources. One terminal acts as the communicator, sending information to each of the remaining m terminals via m different noiseless channels. A communication rate-key rate tradeoff region is identified for this model. Although the model in [18] can be completely solved, it is different from the Csiszár-Narayan model treated in the current paper both in the way that the terminals communicate as well as in the manner in which the eavesdroppers cooperate. In particular, in the model considered in [18], each of the m different links have individual eavesdroppers, but cooperation is not allowed among them. Secrecy is no longer guaranteed if the eavesdroppers cooperate.

Communication complexity has also been studied for two-terminal function computation without a secrecy constraint. In his PhD thesis, Zhao gave an upper bound [19, Theorem 3] on the maximum rate of a *common randomness* achievable from a (non-interactive) communication of a given rate R, the communication being restricted to a single transmission from one terminal to the other. The common randomness here is essentially a function of the source that can be agreed upon by both terminals. There is no requirement that the common randomness be kept secure from an eavesdropper. Courtade [20, Theorem 4] subsequently pointed out a small correction to Zhao's bound in light of the correction made to the strong data processing inequality in [21].

Braverman and Rao in [22, Theorem II.3] gave an exact characterization of the communication complexity for two-party *interactive* function computation.[2] The communication complexity is shown to be equal to an information-theoretic quantity called the *internal information cost*. In a follow-up work, Braverman and Schneider provide an algorithm to compute the internal information cost for binary function computation — see Theorem 1.1 of [23].

Turning our attention to the topic of omnivocality originally studied in our paper [11], Zhang et al. [13] have recently obtained some new results. In particular, their Theorem 5 gives a sufficient condition for when a particular terminal *must* communicate in any SK-capacity-achieving protocol. Our original sufficient condition for omnivocality [11, Theorem 4] (Theorem 10 in this paper) can now be obtained as a consequence of Zhang et al.'s Theorem 5. In addition, Theorem 4 of [13] provides a sufficient condition that guarantees the existence of an SK-capacity-achieving protocol within which a given terminal can remain silent.

C. Organization

The paper is organized as follows. In Section II, we provide the definitions and preliminaries needed for the rest of the paper. In Section III, we state and prove our lower bound on the communication complexity R_{SK}. In Section IV, we identify a class of uniform hypergraph PIN models which are R_{SK}-maximal. Section V identifies a condition that makes omnivocality necessary for achieving SK capacity. The issue of verifying whether that condition holds for a given multi-terminal source model is addressed in Section VI. The effect on R_{SK} of local randomization at the terminals is discussed in Section VII. Finally, Section VIII summarizes our results and presents some open problems. To preserve the flow of the exposition, the proofs of some of our results have been moved to appendices.

[2]To be precise, the quantity which we are calling communication complexity is referred to as *amortized* communication complexity in [22].

II. Preliminaries

We start by giving a mathematical description of the multi-terminal source model of [3]. Throughout, we use $\mathbb{N}$ to denote the set of positive integers. Consider a set of m terminals denoted by $\mathcal{M} = \{1, 2, \ldots, m\}$. Each terminal $i \in \mathcal{M}$ observes n i.i.d. repetitions of the random variable X_i taking values in the finite set $\mathcal{X}_i$. The n i.i.d. copies of the random variable are denoted by X_i^n. For any subset $A \subseteq \mathcal{M}$, X_A and X_A^n denote the collections of random variables $(X_i : i \in A)$ and $(X_i^n : i \in A)$, respectively. The terminals communicate through a noiseless public channel, any communication sent through which is accessible to all terminals and to potential eavesdroppers as well. An *interactive communication* is a communication $\mathbf{f} = (f_1, f_2, \cdots, f_r)$ with finitely many transmissions f_j, in which any transmission sent by the ith terminal is a deterministic function of X_i^n and all the previous communication, i.e., if terminal i transmits f_j, then f_j is a function only of X_i^n and $f_1, \ldots, f_{j-1}$. We denote the random variable associated with $\mathbf{f}$ by $\mathbf{F}$; the support of $\mathbf{F}$ is a finite set $\mathcal{F}$. The rate of the communication $\mathbf{F}$ is defined as $\frac{1}{n}\log|\mathcal{F}|$. Note that $\mathbf{f}$, $\mathbf{F}$ and $\mathcal{F}$ implicitly depend on n.

Definition 1. *A* common randomness (CR) *obtained from an interactive communication $\boldsymbol{F}$ is a sequence of random variables $\boldsymbol{J}^{(n)}$, $n \in \mathbb{N}$, which are functions of $X_{\mathcal{M}}^n$, such that for any $0 < \epsilon < 1$ and for all sufficiently large n, there exist $J_i = J_i(X_i^n, \boldsymbol{F})$, $i = 1, 2, \ldots, m$, satisfying $Pr[J_1 = J_2 = \cdots = J_m = \boldsymbol{J}^{(n)}] \geq 1 - \epsilon$.*

When $\mathbf{J}^{(n)} = X_{\mathcal{M}}^n$ we say that the terminals in $\mathcal{M}$ have attained *omniscience*. The communication $\mathbf{F}$ which achieves this is called a *communication for omniscience*. It was shown in Proposition 1 of [3] that the minimum rate achievable by a communication for omniscience, denoted by R_{CO}, is equal to $\min_{(R_1,R_2,\ldots,R_m)\in\mathcal{R}_{\text{CO}}} \sum_{i=1}^{m} R_i$, where the region $\mathcal{R}_{\text{CO}}$ is given by

$$\mathcal{R}_{\text{CO}} = \left\{ (R_i)_{i\in\mathcal{M}} : \sum_{i\in B} R_i \geq H(X_B|X_{B^c}), B \subsetneq \mathcal{M} \right\}. \quad (1)$$

Henceforth, we will refer to R_{CO} as the "minimum rate of communication for omniscience". Further, it can be seen from the description of $\mathcal{R}_{\text{CO}}$ that $R_{\text{CO}} < \infty$. More precisely, note that the point $(R_1, R_2, \ldots, R_m)$ defined by $R_i = H(X_i)$ for all i lies in $\mathcal{R}_{\text{CO}}$, and hence $R_{\text{CO}} \leq \sum_{i=1}^{m} H(X_i) < \infty$.

Definition 2. *A real number $R \geq 0$ is an* achievable SK rate *if there exists a CR $\boldsymbol{K}^{(n)}$, $n \in \mathbb{N}$, obtained from an interactive communication $\boldsymbol{F}$ satisfying, for any $\epsilon > 0$ and for all sufficiently large n, $I(\boldsymbol{K}^{(n)}; \boldsymbol{F}) \leq \epsilon$ and $\frac{1}{n}H(\boldsymbol{K}^{(n)}) \geq R - \epsilon$. The* SK capacity *is defined to be the supremum among all achievable rates. The CR $\boldsymbol{K}^{(n)}$ is called a* secret key (SK).

From now on, we will drop the superscript (n) from both $\mathbf{J}^{(n)}$ and $\mathbf{K}^{(n)}$ to keep the notation simple.

The SK capacity can be expressed as [3, Theorem 1]

$$\mathcal{C}(\mathcal{M}) = H(X_{\mathcal{M}}) - R_{\text{CO}}. \tag{2}$$

Other equivalent characterizations of $\mathcal{C}(\mathcal{M})$ exist in the literature. Csiszár and Narayan observed in [3, Section V] that since the linear program $R_{\text{CO}} = \min_{(R_1,R_2,\ldots,R_m)\in\mathcal{R}_{\text{CO}}} \sum_{i=1}^{m} R_i$, with the constraints defined by (1), has an optimal solution, by strong duality, the dual linear program also has the same optimal value. Using this fact, the expression for SK capacity can be rewritten as

$$\mathcal{C}(\mathcal{M}) \triangleq H(X_{\mathcal{M}}) - \max_{\lambda\in\Lambda} \sum_{B\in\mathcal{B}} \lambda_B H(X_B|X_{B^c}), \tag{3}$$

where $\mathcal{B}$ is the set of all non-empty, proper subsets of $\mathcal{M}$, and Λ is the set of all *fractional partitions* defined on $\mathcal{B}$. To be precise, any $\lambda = (\lambda_B : B \in \mathcal{B}) \in \Lambda$ satisfies $\lambda_B \geq 0$, for all $B \in \mathcal{B}$, and $\sum_{B:i\in B} \lambda_B = 1$, for all $i \in \mathcal{M}$. It is a fact that $H(X_{\mathcal{M}}) - \max_{\lambda\in\Lambda} \sum_{B\in\mathcal{B}} \lambda_B H(X_B|X_{B^c}) \geq 0$ [24, Proposition II].

Another characterization of SK capacity can be given via the notion of *multipartite information* defined as follows:

$$\mathbf{I}(X_{\mathcal{M}}) \triangleq \min_{\mathcal{P}} \Delta(\mathcal{P}) \tag{4}$$

with $\Delta(\mathcal{P}) \triangleq \frac{1}{|\mathcal{P}|-1}\left[\sum_{A\in\mathcal{P}} H(X_A) - H(X_{\mathcal{M}})\right]$ and the minimum being taken over all partitions $\mathcal{P} = \{A_1, A_2, \cdots, A_\ell\}$ of $\mathcal{M}$, of size $\ell \geq 2$. Note that $\mathbf{I}(X_{\mathcal{M}}^n) = n\mathbf{I}(X_{\mathcal{M}})$. The quantity $\mathbf{I}(X_{\mathcal{M}})$ is a generalization of the mutual information to a multiterminal setting; indeed, for $m = 2$, we have $\mathbf{I}(X_1, X_2) = I(X_1; X_2)$. Note that $\mathbf{I}(X_{\mathcal{M}}) = 0$ iff there exists a partition $\mathcal{P} = \{A_1, A_2, \ldots, A_\ell\}$ of $\mathcal{M}$, with $\ell \geq 2$, such that the random variables $X_{A_1}, X_{A_2}, \ldots, X_{A_\ell}$ are mutually independent. It was shown in Theorem 1.1 of [25] and Theorem 4.1 of [12] that

$$\mathcal{C}(\mathcal{M}) = \mathbf{I}(X_{\mathcal{M}}). \tag{5}$$

For the rest of this paper we shall use $\mathcal{C}(\mathcal{M})$ and $\mathbf{I}(X_{\mathcal{M}})$ interchangeably.

The partition $\{\{1\}, \{2\}, \ldots, \{m\}\}$ consisting of m singleton cells will play a special role in the later sections of this paper; we call this the *singleton partition* and denote it by $\mathcal{S}$. The sources where $\mathcal{S}$ *is a minimizer for* (4) will henceforth be referred to as *Type* $\mathcal{S}$ *sources*. If $\mathcal{S}$ *is the unique minimizer for* (4) then we call such a source *strict Type* $\mathcal{S}$. A connection between the optimal fractional partition in (3) and the optimal partition in (4) was pointed out in [25]. For any partition $\mathcal{P}$ of $\mathcal{M}$ define $\lambda^{(\mathcal{P})}$ as follows: $\lambda_B^{(\mathcal{P})} \triangleq \frac{\mathbb{I}\{B^c\in\mathcal{P}\}}{|\mathcal{P}|-1}$, for all $B \in \mathcal{B}$ and $\mathbb{I}\{.\}$ is the indicator function. It is easy to check that $\lambda^{(\mathcal{P})}$ is a fractional partition on $\mathcal{B}$, and $\Delta(\mathcal{P}) = H(X_{\mathcal{M}}) - \sum_{B\in\mathcal{B}} \lambda_B^{(\mathcal{P})} H(X_B|X_{B^c})$. Hence, for any partition $\mathcal{P}$ which is a minimizer in (4), the corresponding $\lambda^{(\mathcal{P})}$ is an optimal fractional partition for (3).

We are now in a position to make the notion of communication complexity rigorous.

Definition 3. *A real number $R \geq 0$ is said to be an* achievable rate of interactive communication for maximal-rate SK *if for all $\epsilon > 0$ and for all sufficiently large n, there exist* (i) *an interactive communication $\boldsymbol{F}$ satisfying $\frac{1}{n}\log|\mathcal{F}| \leq R+\epsilon$, and* (ii) *an SK $\boldsymbol{K}$ obtained from $\boldsymbol{F}$ such that $\frac{1}{n}H(\boldsymbol{K}) \geq \boldsymbol{I}(X_{\mathcal{M}}) - \epsilon$.*

The infimum among all such achievable rates is called the communication complexity of achieving SK capacity*, denoted by R_{SK}.*

The proof of Theorem 1 in [3] shows that there exists an interactive communication $\mathbf{F}$ that enables omniscience at all terminals and from which a maximal-rate SK can be obtained. Therefore, we have $R_{\text{SK}} \leq R_{\text{CO}} < \infty$. Hence, in terms of communication complexity, the sources that satisfy $R_{\text{SK}} = R_{\text{CO}}$ are the worst-case sources. We will henceforth refer to them as R_{SK}*-maximal* sources. Such sources do exist, as will be shown in Sections IV and VI.

Tyagi gave a characterization of R_{SK} in the case of a two-terminal model [8, Theorem 3].[3] The key to his characterization was the observation that conditioned on a maximal-rate SK $\mathbf{K}$ and the communication $\mathbf{F}$ from which $\mathbf{K}$ is extracted, the observations of the two terminals are "almost" independent: $\frac{1}{n}I(X_1^n; X_2^n|\mathbf{K}, \mathbf{F}) \to 0$ as $n \to \infty$. Thus, the pair $(\mathbf{K}, \mathbf{F})$ is a Wyner common information [10] for the randomness at the terminals. Tyagi used the term "interactive common information" to denote any Wyner common information that consisted of a CR along with the interactive communication achieving it.

We extend Tyagi's ideas to the setting of $m \geq 2$ terminals. We first extend the definition of conditional mutual information to the multiterminal setting. We will refer to the multiterminal analogue of the conditional mutual information as the *conditional multipartite information* (CMI). As a natural extension of (4), we could define CMI as $\min_{\mathcal{P}} \Delta(\mathcal{P}|\mathbf{L})$, where $\Delta(\mathcal{P}|\mathbf{L}) \triangleq \frac{1}{|\mathcal{P}|-1}\left[\sum_{A\in\mathcal{P}} H(X_A|\mathbf{L}) - H(X_{\mathcal{M}}|\mathbf{L})\right]$. Note that

$$\Delta(\mathcal{P}|\mathbf{L}) = H(X_{\mathcal{M}}|\mathbf{L}) - \sum_{B\in\mathcal{B}} \lambda_B^{(\mathcal{P})} H(X_B|X_{B^c}, \mathbf{L}). \tag{6}$$

Using this definition, we can indeed generalize Tyagi's arguments to the case of $m \geq 2$ terminals and obtain a lower bound on R_{SK}. It turns out, however, that a stronger lower bound can be obtained by defining CMI to be equal to $\Delta(\mathcal{P}'|\mathbf{L})$, where $\mathcal{P}'$ is any partition that achieves the minimum in (4). One complication now is that there could be more than one choice of $\mathcal{P}'$ that achieves the minimum in (4), and two distinct choices of $\mathcal{P}'$ could yield different values for $\Delta(\mathcal{P}'|\mathbf{L})$. We simply choose the *finest* partition $\mathcal{P}^*$ that achieves the minimum in (4); we henceforth refer to this partition $\mathcal{P}^*$ as the *fundamental partition* of the source $X_{\mathcal{M}}$. We then define CMI as

$$\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}|\mathbf{L}) \triangleq \Delta(\mathcal{P}^*|\mathbf{L}), \tag{7}$$

[3]It should be clarified that Tyagi's characterization was for "weak" SKs, which are defined as in our Definition 2, except that the security condition $I(\mathbf{K}; \mathbf{F}) \leq \epsilon$ is weakened to $\frac{1}{n}I(\mathbf{K}; \mathbf{F}) \leq \epsilon$. However, using techniques from [26], for example, it may be possible that Tyagi's achievability results can be extended to obtain SKs that satisfy the stronger security condition, without increasing R_{SK}.

the subscript $\mathcal{P}^*$ being used to emphasize the role of the fundamental partition in the definition. It needs to be clarified here that $\mathcal{P}^*$ exists and is unique by Theorem 5.2 of [12]. For Type $\mathcal{S}$ sources it is easy to see that $\mathcal{P}^* = \mathcal{S}$.

The definition of $\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}|\mathbf{L})$ applies to any collection of jointly distributed random variables $X_{\mathcal{M}}$; in particular it applies to the collection $X_{\mathcal{M}}^n$. To be clear,

$$\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\mathbf{L}) \triangleq \frac{1}{|\mathcal{P}^*|-1}\left[\sum_{A\in\mathcal{P}^*} H(X_A^n|\mathbf{L}) - H(X_{\mathcal{M}}^n|\mathbf{L})\right].$$

We point out an important consequence of our definition of $\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}|\mathbf{L})$. We have $\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}|\mathbf{L}) = 0$ iff the random variables $X_{\mathcal{M}}$ are conditionally independent across the cells of $\mathcal{P}^*$ given $\mathbf{L}$. We are now in a position to extend the notion of the Wyner common information to a multipartite setting.[4]

Definition 4. *A* (multiterminal) Wyner common information (CI_W) *for $X_{\mathcal{M}}$ is a sequence of finite-valued functions $\boldsymbol{L}^{(n)} = \boldsymbol{L}^{(n)}(X_{\mathcal{M}}^n)$ such that $\frac{1}{n}\boldsymbol{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\boldsymbol{L}^{(n)}) \to 0$ as $n \to \infty$. An* interactive common information (CI) *for $X_{\mathcal{M}}$ is a Wyner common information of the form $\boldsymbol{L}^{(n)} = (\boldsymbol{J}, \boldsymbol{F})$, where $\boldsymbol{F}$ is an interactive communication and $\boldsymbol{J}$ is a CR obtained from $\boldsymbol{F}$.*

Again, we shall drop the superscript (n) from $\mathbf{L}^{(n)}$ for notational simplicity. Wyner common informations $\mathbf{L}$ do exist: for example, the identity map $\mathbf{L} = X_{\mathcal{M}}^n$ is a CI_W. To see that CIs $(\mathbf{J}, \mathbf{F})$ also exist, observe that $\mathbf{J} = X_{\mathcal{M}}^n$ and a communication $\mathbf{F}$ enabling omniscience constitute a CI_W, and hence, a CI.

Definition 5. *A real number $R \geq 0$ is an* achievable CI_W (resp. CI) rate *if there exists a CI_W $\boldsymbol{L}$ (resp. a CI $\boldsymbol{L} = (\boldsymbol{J}, \boldsymbol{F})$) such that for all $\epsilon > 0$, we have $\frac{1}{n}H(\boldsymbol{L}) \leq R + \epsilon$ for all sufficiently large n. We denote the infimum among all achievable CI_W (resp. CI) rates by $\text{CI}_W(X_{\mathcal{M}})$ (resp. $\text{CI}(X_{\mathcal{M}})$).*

The proposition below records the relationships between some of the information-theoretic quantities defined so far.

Proposition 1. *For a multiterminal source $X_{\mathcal{M}}^n$, we have $H(X_{\mathcal{M}}) \geq \text{CI}(X_{\mathcal{M}}) \geq \text{CI}_W(X_{\mathcal{M}}) \geq \boldsymbol{I}(X_{\mathcal{M}})$.*

Proof: The first inequality is due to the fact that there exists a CI of rate $H(X_{\mathcal{M}})$. The second follows from the fact that a CI is a special type of CI_W, so that $\text{CI}(X_{\mathcal{M}}) \geq \text{CI}_W(X_{\mathcal{M}})$.

For the last inequality, we start by observing that for any function $\mathbf{L}$ of $X_{\mathcal{M}}^n$, we have

$$\begin{aligned}\mathbf{I}(X_{\mathcal{M}}^n) &- \mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\mathbf{L}) \\ &= I(X_{\mathcal{M}}^n; \mathbf{L}) - \sum_{B\in\mathcal{B}} \lambda_B^{(\mathcal{P}^*)} I(X_B^n; \mathbf{L}|X_{B^c}^n) \qquad (8)\\ &= H(\mathbf{L}) - \sum_{B\in\mathcal{B}} \lambda_B^{(\mathcal{P}^*)} H(\mathbf{L}|X_{B^c}^n), \qquad (9)\end{aligned}$$

where (8) follows from (6) and (7) and hence,

$$\frac{1}{n}H(\mathbf{L}) = \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\mathbf{L}). \qquad (10)$$

[4]In fact, there exist other ways of generalizing Wyner common information to the multiterminal setting (see [27] and [28]).

Now, if $\mathbf{L}$ is any CI_W of rate R, then by Definitions 4 and 5, for every $\epsilon > 0$, we have $\frac{1}{n}H(\mathbf{L}) \leq R+\epsilon$ and $\frac{1}{n}\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\mathbf{L}) \leq \epsilon$ for all sufficiently large n. Thus, in conjunction with (10), we have $R + \epsilon \geq \frac{1}{n}H(\mathbf{L}) \geq \mathbf{I}(X_{\mathcal{M}}) - \epsilon$ for all sufficiently large n. In particular, $R + \epsilon \geq \mathbf{I}(X_{\mathcal{M}}) - \epsilon$ holds for any $\epsilon > 0$, from which we infer that $R \geq \mathbf{I}(X_{\mathcal{M}})$. The inequality $\text{CI}_W(X_{\mathcal{M}}) \geq \mathbf{I}(X_{\mathcal{M}})$ now follows. ∎

Finally, analogous to Definition 3, we have a definition of achievable rate of interactive communication required to get a CI.

Definition 6. *A real number $R \geq 0$ is said to be an* achievable rate of interactive communication for CI *if for all $\epsilon > 0$ and for all sufficiently large n, there exist* (i) *an interactive communication $\boldsymbol{F}$ satisfying $\frac{1}{n}\log|\mathcal{F}| \leq R + \epsilon$, and* (ii) *a CR $\boldsymbol{J}$ such that $\boldsymbol{L} = (\boldsymbol{J}, \boldsymbol{F})$ is a CI. We denote the infimum among all such achievable rates by R_{CI}.*

III. Lower Bound on R_{SK}

The goal of this section is to state and prove a lower bound on R_{SK}, which partially extends Tyagi's two-terminal result [8, Theorem 3] to the multiterminal setting.

Theorem 2. *For any multiterminal source $X_{\mathcal{M}}$, we have*

$$R_{SK} \geq R_{CI} \geq \text{CI}(X_{\mathcal{M}}) - \boldsymbol{I}(X_{\mathcal{M}}).$$

By Proposition 1, the lower bounds above are non-negative.

The ideas in our proof of Theorem 2 may be viewed as a natural extension of those in the proof of [8, Theorem 3]. We start with three preliminary lemmas. In all that follows, $\lambda^* = \lambda^{(\mathcal{P}^*)}$, and moreover, $\lambda^* = (\lambda_B^* : B \in \mathcal{B})$. The first lemma is the multiterminal extension of [8, Lemma 7].

Lemma 3. *For any function $\boldsymbol{L}$ of $X_{\mathcal{M}}$, we have*

$$n\boldsymbol{I}(X_{\mathcal{M}}) = \boldsymbol{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\boldsymbol{L}) + H(\boldsymbol{L}) - \sum_{B\in\mathcal{B}} \lambda_B^* H(\boldsymbol{L}|X_{B^c}^n).$$

Proof: The result is just a restatement of (9). ∎

Lemma 4. *For any CR $\boldsymbol{J}$ obtained from an interactive communication $\boldsymbol{F}$,*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{B\in\mathcal{B}} \lambda_B^* H(\boldsymbol{J}|X_{B^c}^n, \boldsymbol{F}) = 0.$$

Proof: Fix an $\epsilon > 0$. We have for all sufficiently large n, by Fano's inequality,

$$\begin{aligned}\frac{1}{n}\sum_{B\in\mathcal{B}} \lambda_B^* H(\mathbf{J}|X_{B^c}^n, \mathbf{F}) &\leq \frac{1}{n}\sum_{B\in\mathcal{B}} \lambda_B^* \left(h(\epsilon) + \epsilon H(X_{B^c}^n, \mathbf{F})\right)\\ &\leq \frac{1}{n}\sum_{B\in\mathcal{B}} \lambda_B^* \left(h(\epsilon) + \epsilon H(X_{\mathcal{M}}^n, \mathbf{F})\right)\\ &= \frac{1}{n}\sum_{B\in\mathcal{B}} \lambda_B^* \left(h(\epsilon) + \epsilon H(X_{\mathcal{M}}^n)\right)\\ &= \frac{1}{n}\sum_{B\in\mathcal{B}} \lambda_B^* \left(h(\epsilon) + n\epsilon H(X_{\mathcal{M}})\right)\end{aligned}$$

$$\leq (2^m - 2)\left[h(\epsilon) + \epsilon H(X_{\mathcal{M}})\right], \tag{11}$$

where $h(.)$ is the binary entropy function, and (11) follows from the fact that, by definition, $\lambda_B^* \leq 1$ and $|\mathcal{B}| = 2^m - 2$. Note that the expression in (11) goes to 0 with ϵ, since $h(\epsilon) \to 0$ as $\epsilon \to 0$, and $H(X_{\mathcal{M}}) \leq \log(\prod_{j=1}^{m} |\mathcal{X}_j|)$. ∎

The last lemma we need, stated without proof, is a special case of [4, Lemma B.1]. It is the multiterminal extension of [8, Lemma 5].

Lemma 5 ([4], Lemma B.1)**.** *For an interactive communication $\boldsymbol{F}$ we have*

$$H(\boldsymbol{F}) \geq \sum_{B \in \mathcal{B}} \lambda_B^* H(\boldsymbol{F} | X_{B^c}^n).$$

With these lemmas in hand, we can proceed to the proof of Theorem 2.

Proof of Theorem 2: The proof is done in two parts. In the first part, we prove that $R_{\text{CI}} \geq \text{CI}(X_{\mathcal{M}}) - \mathbf{I}(X_{\mathcal{M}})$. In the second part, we show that $R_{\text{SK}} \geq R_{\text{CI}}$.

Part I: $R_{CI} \geq CI(X_{\mathcal{M}}) - \boldsymbol{I}(X_{\mathcal{M}})$

The idea is to show that $\mathbf{I}(X_{\mathcal{M}}) + R_{\text{CI}}$ is an achievable CI rate, so that $\text{CI}(X_{\mathcal{M}}) \leq \mathbf{I}(X_{\mathcal{M}}) + R_{\text{CI}}$.

Fix an $\epsilon > 0$. By the definition of R_{CI}, for all sufficiently large n, there exists an interactive communication $\mathbf{F}$ satisfying $\frac{1}{n}\log|\mathcal{F}| \leq R_{\text{CI}} + \epsilon/2$ and a CR $\mathbf{J}$ such that $\mathbf{L} = (\mathbf{J}, \mathbf{F})$ is a CI. We will show that $\frac{1}{n}H(\mathbf{J}, \mathbf{F}) \leq \mathbf{I}(X_{\mathcal{M}}) + R_{\text{CI}} + \epsilon$ for all sufficiently large n. This, by Definition 5, shows that $\mathbf{I}(X_{\mathcal{M}}) + R_{\text{CI}}$ is an achievable CI rate.

Setting $\mathbf{L} = (\mathbf{J}, \mathbf{F})$ in Lemma 3, we obtain

$$\begin{aligned}
&\frac{1}{n}\left[H(\mathbf{J},\mathbf{F}) - \sum_{B \in \mathcal{B}} \lambda_B^* H(\mathbf{F}|X_{B^c}^n)\right] - \mathbf{I}(X_{\mathcal{M}}) \\
&\quad = \frac{1}{n}\left[\sum_{B \in \mathcal{B}} \lambda_B^* H(\mathbf{J}|X_{B^c}^n, \mathbf{F}) - \mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\mathbf{J}, \mathbf{F})\right] \\
&\quad \leq \epsilon/2, \qquad (12)
\end{aligned}$$

where (12) follows from Lemma 4. Re-arranging, we get

$$\begin{aligned}
\frac{1}{n}H(\mathbf{J}, \mathbf{F}) &\leq \mathbf{I}(X_{\mathcal{M}}) + \frac{1}{n}\sum_{B \in \mathcal{B}} \lambda_B^* H(\mathbf{F}|X_{B^c}^n) + \epsilon/2 \\
&\leq \mathbf{I}(X_{\mathcal{M}}) + \frac{1}{n}H(\mathbf{F}) + \epsilon/2,
\end{aligned}$$

the second inequality coming from Lemma 5. Finally, using the fact that $\frac{1}{n}H(\mathbf{F}) \leq \frac{1}{n}\log|\mathcal{F}| \leq R_{\text{CI}} + \epsilon/2$, we see that

$$\frac{1}{n}H(\mathbf{J}, \mathbf{F}) \leq \mathbf{I}(X_{\mathcal{M}}) + R_{\text{CI}} + \epsilon$$

which is what we set out to prove.

Part II: $R_{SK} \geq R_{CI}$

Fix $\epsilon > 0$. From the definition of R_{SK}, there exist an interactive communication $\mathbf{F}$ and an SK $\mathbf{K}$ obtained from $\mathbf{F}$ such that, for all sufficiently large n, $\frac{1}{n}\log|\mathcal{F}| \leq R_{SK} + \epsilon$ and $\frac{1}{n}H(\mathbf{K}) \geq \mathbf{I}(X_{\mathcal{M}}) - \epsilon$. We wish to show that $(\mathbf{K}, \mathbf{F})$ is a CI, so that by Definition 6, we would have $R_{\text{SK}} \geq R_{\text{CI}}$.

Setting $\mathbf{L} = (\mathbf{K}, \mathbf{F})$ in Lemma 3, we have for all sufficiently large n,

$$\begin{aligned}
&\frac{1}{n}\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|\mathbf{K}, \mathbf{F}) \\
&\quad = \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}, \mathbf{F}) + \frac{1}{n}\sum_{B \in \mathcal{B}} \lambda_B^* H(\mathbf{F}|X_{B^c}^n) \\
&\qquad + \frac{1}{n}\sum_{B \in \mathcal{B}} \lambda_B^* H(\mathbf{K}|X_{B^c}^n, \mathbf{F}) \\
&\quad \leq \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}|\mathbf{F}) + \epsilon \qquad (13) \\
&\quad \leq \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}) + \epsilon + \epsilon \qquad (14) \\
&\quad \leq 3\epsilon, \qquad (15)
\end{aligned}$$

where (13) follows from Lemmas 4 and 5, (14) follows from the fact that $I(\mathbf{K}; \mathbf{F}) \leq \epsilon$, while (15) is due to the fact that $\frac{1}{n}H(\mathbf{K}) \geq \mathbf{I}(X_{\mathcal{M}}) - \epsilon$. Thus, by Definition 4, $(\mathbf{K}, \mathbf{F})$ is a CI. ∎

An issue with our Theorem 2 is that the bounds are difficult to evaluate explicitly, as we do not have a computable characterization of $\text{CI}(X_{\mathcal{M}})$. For the special case of the two-terminal model, Theorem 3 of [8] shows that both the bounds in Theorem 2 are tight. However, at least one of the bounds in our Theorem 2 is not tight in general, as shown by the following example.

Example III.1. *Let Z_1^n, Z_2^n, Z_3^n and Z_4^n be four independent i.i.d. sequences of Bernoulli$(1/2)$ random variables, and let $\mathcal{M} = \{1, 2, 3, 4\}$. We construct a source $X_{\mathcal{M}}^n$ as follows:*

$$\begin{aligned}
X_1^n &= (Z_1^n, Z_2^n), \\
X_2^n &= (Z_1^n, Z_3^n), \\
X_3^n &= (Z_2^n, Z_3^n, Z_4^n), \text{ and} \\
X_4^n &= Z_4^n.
\end{aligned}$$

It is straightforward to verify that $\mathcal{P}^* = \{\{1, 2, 3\}, \{4\}\}$ and $\boldsymbol{I}(X_{\mathcal{M}}) = 1$.

Now, consider $\boldsymbol{J} = Z_4^n$, and let $\boldsymbol{F}$ be a communication where terminal 3 broadcasts Z_4^n and every other terminal remains silent. Thus, $\boldsymbol{J}$ is a CR achievable with $\boldsymbol{F}$. Also, $H(X_1^n, X_2^n, X_3^n|\boldsymbol{J}, \boldsymbol{F}) + H(X_4^n|\boldsymbol{J}, \boldsymbol{F}) - H(X_{\mathcal{M}}^n|\boldsymbol{J}, \boldsymbol{F}) = 0$, so that $(\boldsymbol{J}, \boldsymbol{F})$ is in fact a CI. Therefore, via Definition 5, we have $CI(X_{\mathcal{M}}) \leq \frac{1}{n}H(\boldsymbol{J}, \boldsymbol{F}) = 1$, and hence by Proposition 1, again noting that $\boldsymbol{I}(X_{\mathcal{M}}) = 1$, we have $CI(X_{\mathcal{M}}) - \boldsymbol{I}(X_{\mathcal{M}}) = 0$. Theorem 2 thus gives the lower bound $R_{SK} \geq 0$. However, since the combined observations of terminals 1 and 2 are independent of the observation of terminal 4, these terminals cannot agree upon an SK of rate 1 without a non-zero rate of communication. This shows that at least one of the inequalities in Theorem 2 is not tight.

In the general multiterminal model, with $m \geq 3$, the best known upper bound on R_{SK} is the minimum rate of communication for omniscience, R_{CO}. In the following section, we identify a large class of sources where our lower bound equals R_{CO}, i.e., the sources are R_{SK}-maximal.

IV. R_{SK}-MAXIMALITY IN UNIFORM HYPERGRAPH PIN MODELS

This section focuses on a special class of sources called the PIN model, introduced in [5] and [6]. A broad class of PIN models defined on uniform hypergraphs (which is a generalization of the PIN models of [5] and [6]) is identified to be R_{SK}-maximal in this section.

In our view, PIN models form a natural class of models for SK agreement, wherein there are various independent SKs being initially shared within various small subsets of terminals, and these SKs need to be combined and distilled to form a group key common to all terminals. The original PIN model on graphs, as defined in [5], was motivated by multipath wireless networks, where the channels between pairs of terminals are reciprocal, and uncorrelated across different pairs of terminals and across different time-coherence intervals. The PIN model we study in this paper is due to [6], which is a specialized version of the earlier PIN model of [5]. PIN models on hypergraphs have been studied extensively in the context of the CCDE problem [14],[16],[17], where subsets of terminals share a common random variable, independent of those shared between other subsets.

Briefly, a *hypergraph PIN model*[5] is defined on an underlying hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \mathcal{M}$, the set of m terminals of the model, and $\mathcal{E}$ being a *multiset* of hyperedges, i.e., subsets of $\mathcal{V}$.[6] For a hyperedge e having ℓ copies in the multiset $\mathcal{E}$, we represent the different copies as $e_1, e_2, \ldots, e_\ell$. To keep the notation simple, if a hyperedge e has only one copy in $\mathcal{E}$, we simply represent it as e instead of e_1. Unless otherwise stated, we will assume that each hyperedge in $\mathcal{E}$ has only one copy. For $n \in \mathbb{N}$, we define $\mathcal{E}^{(n)}$ to be the multiset of hyperedges formed by taking n copies of each element of the multiset $\mathcal{E}$. Associated with each hyperedge $e \in \mathcal{E}^{(n)}$ is a Bernoulli$(1/2)$ random variable ξ_e; the ξ_es are all mutually independent. With this, the random variables X_i^n, $i \in \mathcal{M}$, are defined as $X_i^n = (\xi_e : e \in \mathcal{E}^{(n)} \text{ and } i \in e)$. Each random variable ξ_e, $e \in \mathcal{E}^{(n)}$, should be thought of as one bit of SK initially shared among the terminals in e. By allowing $\mathcal{E}$ to be a multiset in our model, so that a hyperedge can have multiple copies in $\mathcal{E}$, we allow the initial number of bits of SK shared within one subset of terminals (i.e., one hyperedge) to be different from that shared within another subset (hyperedge) — see Appendix D for an example.

When every $e \in \mathcal{E}$ satisfies $|e| = t$, we call $\mathcal{H}$ a *t-uniform hypergraph*. We will show that any Type $\mathcal{S}$ uniform hypergraph PIN model is R_{SK}-maximal.

Theorem 6. *For a Type $\mathcal{S}$ PIN model defined on an underlying t-uniform hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$, we have $CI(X_\mathcal{M}) = CI_W(X_\mathcal{M}) = H(X_\mathcal{M})$, and hence, $R_{SK} = R_{CI} = R_{CO} = \frac{m-t}{m-1}|\mathcal{E}|$.*

[5]Strictly speaking, the term "hypergraph PIN model" is a misnomer, as this model does not in general have the pairwise independence property that characterizes the PIN models on graphs considered in [6]. We have simply chosen to retain the "PIN model" nomenclature for backward compatibility with [6].

[6]Note that we allow $\mathcal{E}$ to contain multiple copies of a hyperedge. In the graph theory literature, such a hypergraph is sometimes referred to as a "multi-hypergraph".

Type $\mathcal{S}$ PIN models defined on t-uniform hypergraphs do indeed exist, as we will see in Section VI. Also, it is possible to efficiently determine if a given source $\mathcal{X}_\mathcal{M}$ (not necessarily a PIN model) is Type $\mathcal{S}$; a strongly polynomial-time algorithm for this has been given by Chan et al. [12]. In Section VI, we present another useful, but inefficient, test for deciding the Type $\mathcal{S}$ property.

The proof of Theorem 6 will require a technical lemma which we state below.

Lemma 7. *For any t-*uniform hypergraph *PIN model and any function $\boldsymbol{L}$ of $X_\mathcal{M}^n$ we have*

$$\sum_{i=1}^{m} I(X_i^n; \boldsymbol{L}) \leq tH(\boldsymbol{L}). \tag{16}$$

The lengthy proof of this lemma is deferred to Appendix A.

Proof of Theorem 6: Observe that $\lambda_B^{(\mathcal{S})} = \frac{1}{m-1}$, whenever $|B| = m-1$ and $\lambda_B^{(\mathcal{S})} = 0$, otherwise. Hence, for any Type $\mathcal{S}$ source $X_\mathcal{M}^n$, we have

$$\mathbf{I}_{\mathcal{P}^*}(X_\mathcal{M}^n|\mathbf{L}) = H(X_\mathcal{M}^n|\mathbf{L}) - \frac{1}{m-1}\sum_{i=1}^{m} H(X_{\mathcal{M}\setminus\{i\}}^n|X_i^n, \mathbf{L}) \tag{17}$$

using (6) and (7). Now assume that $X_\mathcal{M}$ arises from a PIN model defined on a t-uniform hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$, and consider any function $\mathbf{L}$ of $X_\mathcal{M}^n$. This allows us to further simplify (17):

$$\begin{aligned}
&\mathbf{I}_{\mathcal{P}^*}(X_\mathcal{M}^n|\mathbf{L}) \\
&\quad= H(X_\mathcal{M}^n) - H(\mathbf{L}) \\
&\qquad - \frac{1}{m-1}\sum_{i=1}^{m}\left[H(X_\mathcal{M}^n) - H(X_i^n) - H(\mathbf{L}|X_i^n)\right] \\
&\quad= \frac{n(t-1)|\mathcal{E}|}{m-1} - H(\mathbf{L}) + \frac{1}{m-1}\sum_{i=1}^{m} H(\mathbf{L}|X_i^n) \qquad (18)\\
&\quad= \frac{n(t-1)|\mathcal{E}|}{m-1} - \frac{1}{m-1}\left[\sum_{i=1}^{m} I(X_i^n;\mathbf{L}) - H(\mathbf{L})\right] \\
&\quad= \frac{n(t-1)}{m-1}\left(|\mathcal{E}| - \frac{1}{n}H(\mathbf{L})\right) \\
&\qquad - \frac{1}{m-1}\left[\sum_{i=1}^{m} I(X_i^n;\mathbf{L}) - tH(\mathbf{L})\right] \\
&\quad\geq \frac{n(t-1)}{m-1}\left(|\mathcal{E}| - \frac{1}{n}H(\mathbf{L})\right), \qquad (19)
\end{aligned}$$

the equality (18) using the facts that $H(X_\mathcal{M}^n) = n|\mathcal{E}|$ and $\sum_{i=1}^m H(X_i^n) = nt|\mathcal{E}|$, and (19) following from Lemma 7.

We will now compute $\text{CI}(X_\mathcal{M})$ using Proposition 1. The upper bound gives us $\text{CI}(X_\mathcal{M}) \leq |\mathcal{E}|$, as $H(X_\mathcal{M}) = |\mathcal{E}|$. For the lower bound, let $\mathbf{L}$ be any CI_W so that for any $\epsilon > 0$, we have $\frac{1}{n}\mathbf{I}_{\mathcal{P}^*}(X_\mathcal{M}^n|\mathbf{L}) < \frac{(t-1)\epsilon}{(m-1)}$ for all sufficiently large n. The bound in (19) thus yields $\frac{1}{n}H(\mathbf{L}) > |\mathcal{E}| - \epsilon$ for all sufficiently large n. Hence, it follows that $\text{CI}_W(X_\mathcal{M}) \geq |\mathcal{E}|$. From the upper and lower bounds in Proposition 1, we then obtain $CI_W(X_\mathcal{M}) = CI(X_\mathcal{M}) = H(X_\mathcal{M})$.

Now from Theorem 2 we have $R_{\text{SK}} \geq R_{\text{CI}} \geq CI(X_\mathcal{M}) -$

$\mathbf{I}(X_{\mathcal{M}})$. Hence, we have

$$R_{\text{SK}} \geq R_{\text{CI}} \geq |\mathcal{E}| - \mathbf{I}(X_{\mathcal{M}}) = H(X_{\mathcal{M}}) - \mathbf{I}(X_{\mathcal{M}}) = R_{\text{CO}}, \tag{20}$$

where the last equality is from (2). But we also have $R_{\text{SK}} \leq R_{\text{CO}}$, as pointed out in Section II, which proves that $R_{\text{SK}} = R_{\text{CI}} = R_{\text{CO}}$.

To obtain the exact expression for R_{CO}, we note that by (2) and (4), $R_{\text{CO}} = H(X_{\mathcal{M}}) - \Delta(\mathcal{S}) = \frac{m}{m-1} H(X_{\mathcal{M}}) - \frac{1}{m-1} \sum_{i=1}^{m} H(X_i)$. This simplifies to the expression stated in the theorem using the facts (already mentioned above) that $H(X_{\mathcal{M}}) = |\mathcal{E}|$ and $\sum_{i=1}^{m} H(X_i) = t|\mathcal{E}|$. ∎

It turns out that for PIN models on graphs (i.e., $t = 2$), the Type $\mathcal{S}$ condition is also necessary for R_{SK}-maximality. It is possible that this holds for PIN models on t-uniform hypergaphs (with $t \geq 3$) as well, but we do not have a proof for this yet.

Theorem 8. *A PIN model defined on a graph is R_{SK}-maximal iff it is Type $\mathcal{S}$.*

We will prove the necessity of the Type $\mathcal{S}$ condition by showing that any graph PIN model that is not Type $\mathcal{S}$ has an SK-capacity-achieving protocol of communication rate strictly less than R_{CO}. To do this, we need a few preliminaries. Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and define $\mathcal{G}^{(n)} = (\mathcal{V}, \mathcal{E}^{(n)})$ for any positive integer n. The *spanning tree packing number* of $\mathcal{G}^{(n)}$, denoted by $\sigma(\mathcal{G}^{(n)})$, is the maximum number of edge-disjoint spanning trees of $\mathcal{G}^{(n)}$. It is a fact that $\lim_{n\to\infty} \frac{1}{n} \sigma(G^{(n)})$ exists (see [6, Proposition 4]); we denote this limit by $\overline{\sigma}(\mathcal{G})$ and call it the *spanning tree packing rate* of the graph $\mathcal{G}$. It was shown in [6, Theorem 5] that for a PIN model on $\mathcal{G}$, we have $\mathcal{C}(\mathcal{M}) = \overline{\sigma}(\mathcal{G})$. Therefore, by (2) we have $R_{\text{CO}} = H(X_{\mathcal{M}}) - \overline{\sigma}(\mathcal{G}) = |\mathcal{E}| - \overline{\sigma}(\mathcal{G})$. We also have the following lemma, the proof of which is given in Appendix B.

Lemma 9. *For a PIN model defined on a graph $\mathcal{G}$, we have $R_{SK} \leq (m-2)\overline{\sigma}(\mathcal{G})$.*

Proof of Theorem 8: The "if" part follows from Theorem 6. For the "only if" part, consider a PIN model on $\mathcal{G}$ that is not of Type $\mathcal{S}$. Using the fact that SK capacity equals the spanning tree packing rate, we then have via (4)

$$\overline{\sigma}(\mathcal{G}) = \mathcal{C}(\mathcal{M}) < \Delta(\mathcal{S}) = \frac{|\mathcal{E}|}{m-1}.$$

Therefore, $|\mathcal{E}| > (m-1)\overline{\sigma}(\mathcal{G})$, or equivalently, $|\mathcal{E}| - \overline{\sigma}(\mathcal{G}) > (m-2)\overline{\sigma}(\mathcal{G})$. Since $R_{\text{CO}} = |\mathcal{E}| - \overline{\sigma}(\mathcal{G})$, we obtain $R_{\text{CO}} > R_{\text{SK}}$ via Lemma 9. ∎

It is natural to ask at this point whether all Type $\mathcal{S}$ sources (not necessarily PIN models) are R_{SK}-maximal. The answer turns out to be "No", as shown by the following example.

Example IV.1. *Let W be a Ber(p) rv, for some $p \in [0,1]$: $\Pr[W=1] = 1 - \Pr[W=0] = p$. Let $X_1, \ldots, X_m$ be random variables that are conditionally independent given W, with*

$$\Pr[X_i = 01 | W = 0] = 1 - \Pr[X_i = 00 | W = 0] = 0.5$$

and

$$\Pr[X_i = 11 | W = 1] = 1 - \Pr[X_i = 10 | W = 1] = 0.5$$

for $i = 1, 2, \ldots, m$. Denote by $h(p)$ the binary entropy of p.

It is easy to check that $H(X_A) = |A| + h(p)$ for all $A \subseteq \mathcal{M}$, and $H(X_i|X_j) = 1$ for all distinct $i, j \in \mathcal{M}$. Therefore, all partitions $\mathcal{P}$ of $\mathcal{M}$ satisfy $\Delta(\mathcal{P}) = h(p)$, and hence, $\boldsymbol{I}(X_{\mathcal{M}}) = h(p)$. In particular, $X_{\mathcal{M}}$ defines a Type $\mathcal{S}$ source. Furthermore, using (2)*, we have $R_{CO} = m$.*

We now show that $R_{SK} < R_{CO}$. Consider a Slepian-Wolf code (see [29, Section 10.3.2]) of rate $H(X_1|X_2) = 1$ for terminal 1. All terminals can recover X_1^n since $H(X_1|X_i) = 1$ for all $i \in \{2, 3, \cdots, m\}$. Then, using the balanced coloring lemma [3, Lemma B3] on X_1^n, an SK of rate $H(X_1) - H(X_1|X_2) = h(p)$ can be obtained. Hence, $R_{SK} \leq 1 < m = R_{CO}$.

In fact, there exist non R_{SK}-maximal sources with $\mathcal{S}$ being a *unique* minimizer for (4). We provide one such example in Appendix C.

V. Omnivocality: When is it necessary?

It is a well-established fact (see [1], [2]) that to generate a maximal-rate SK within a two-terminal source model, it is enough for only one terminal to communicate.[7] So it is natural to ask whether this fact extends to the general multiterminal setting. In other words, for $m \geq 3$, is there always an SK generation protocol involving $m-1$ or fewer terminals communicating that achieves SK capacity? If not, can we identify a class of sources where omnivocality, i.e., all terminals communicating, is required to achieve SK capacity? This section addresses these questions. The main result of this section says that if a source is strict Type $\mathcal{S}$, then omnivocality is required for achieving SK capacity.

Theorem 10. *For a strict Type $\mathcal{S}$ source on $m \geq 3$ terminals, omnivocal communication is necessary for achieving SK capacity.*

There indeed exist sources which are strict Type $\mathcal{S}$. We give a few examples of such sources in Section VI.

Theorem 10 gives a sufficient condition for identifying sources where omnivocality is necessary to generate a maximal-rate SK. The next result shows that the condition is also necessary when $m = 3$, i.e., for any source on 3 terminals which is not strict Type $\mathcal{S}$, there always exists a non-omnivocal key generation protocol that leads to SK capacity.

Theorem 11. *In the three-terminal source model, omnivocal communication is necessary for achieving SK capacity iff the singleton partition $\mathcal{S}$ is the unique minimizer for $\boldsymbol{I}(X_{\mathcal{M}})$ in* (4)*.*

[7] To be precise, the results in [1] and [2] are based on a weaker notion of secrecy, where in Definition 2, the condition $I(\mathbf{K}; \mathbf{F}) \leq \epsilon$ is replaced by $\frac{1}{n} I(\mathbf{K}; \mathbf{F}) \leq \epsilon$. However, it can be shown that one terminal communicating suffices to achieve SK capacity for $m = 2$, in the stronger sense as in Definition 2. Terminal 1 uses a Slepian-Wolf code of rate $H(X_1|X_2)$ to communicate X_1^n to terminal 2. Both terminals now use a balanced coloring (see [3, Lemma B.3]) on X_1^n to get a strong SK of rate $I(X_1; X_2) = \mathbf{I}(X_1, X_2)$.

A conjecture was made in [11] that the necessity of omnivocality implies a strict Type $\mathcal{S}$ source for any $m \geq 3$. It turns out that the conjecture is incorrect. Chan et al. have found an explicit example [12, extension to Example C.1] of a non-strict Type $\mathcal{S}$ PIN model on $m > 3$ terminals that requires omnivocality to achieve SK capacity. For ease of reference, we reproduce this example in Appendix D.

The recent work of Zhang et al. [13] also deserves a special mention here. This work addresses the somewhat more general question of when communication from a specific terminal $i \in \mathcal{M}$ is necessary in order to achieve SK capacity. Theorem 4 of Zhang et al. gives a sufficient condition for the existence of an SK-capacity-achieving protocol in which terminal i remains silent, while their Theorem 5 gives conditions under which terminal i must talk in order to achieve SK capacity. Indeed, the sufficient condition for omnivocality we present in our Theorem 10 follows easily from their Theorem 5. As pointed out by Zhang et al., their theorems almost completely resolve the question they address, except for one case that is not covered by the theorems. They conjectured that in this case, SK capacity could be achieved with terminal i remaining silent. Footnote 4 of [12] reports that using the successive omniscience techniques developed in [30], the conjecture of Zhang et al. can be proved, thereby solving the problem completely.

We now turn to the proofs of Theorems 10 and 11. We prove the former theorem first. The main technical result used in the proof is the SK capacity with silent terminals by Gohari and Anantharam in [7, Theorem 6]. More precisely, suppose we restrict ourselves to SK generation protocols where, only an arbitrary subset of terminals $T \subset \mathcal{M}$ is allowed to communicate. We denote the maximum rate of SK that can be generated by such protocols by $\mathbf{I}_T(X_\mathcal{M})$. Then we have[8]

Theorem 12 (Theorem 6, [7])**.** *For any $T \subset \mathcal{M}$, $\boldsymbol{I}_T(X_\mathcal{M}) = H(X_T) - R_T^{(\min)}$, where $R_T^{(\min)} = \min_{\mathsf{R} \in \mathcal{R}_T} \sum_{i \in T} R_i$, with*

$$\mathcal{R}_T = \Bigg\{ \mathsf{R} = (R_i, i \in T) : \sum_{i \in B \cap T} R_i \geq H(X_{B \cap T} | X_{B^c}), \quad \forall\, B \subsetneq \mathcal{M},\ B \cap T \neq \emptyset \Bigg\}. \tag{21}$$

Note that if $\mathbf{I}(X_\mathcal{M}) > \mathbf{I}_T(X_\mathcal{M})$ for all $T \subset \mathcal{M}$ of size $|T| = m - 1$, then omnivocality is necessary for achieving SK capacity. Thus, our approach for showing that omnivocal communication is needed in certain cases is to use Theorem 12 to prove that $\mathbf{I}(X_\mathcal{M}) > \mathbf{I}_T(X_\mathcal{M})$ for all $(m-1)$-subsets $T \subset \mathcal{M}$. For this, we will need a lower bound on $R_T^{(\min)}$ when $|T| = m - 1$. To prove this bound, we make use of a simpler characterization (than that given in Theorem 12) of the rate region $\mathcal{R}_T$ when $|T| = m - 1$.

Lemma 13. *Let $T = \mathcal{M} \setminus \{u\}$ for some $u \in \mathcal{M}$. The rate region $\mathcal{R}_T$ is the set of all points $(R_i,\, i \in T)$ such that*

$$\sum_{i \in B} R_i \geq H(X_B | X_{T \setminus B}) \quad \forall\, B \subsetneq T, B \neq \emptyset, \text{ and} \qquad \sum_{i \in T} R_i \geq H(X_T | X_u). \tag{22}$$

Proof: Observe that $\mathcal{R}_T$ is defined by constraints on sums of the form $\sum_{i \in B'} R_i$ for non-empty subsets $B' \subseteq T$. When $B' = T$, the constraint is simply $\sum_{i \in T} R_i \geq H(X_T | X_u)$.

Now, consider any non-empty $B' \subsetneq T$. From Theorem 12, we see that constraints on $\sum_{i \in B'} R_i$ arise as constraints on $\sum_{i \in B \cap T} R_i$ in two ways: when $B = B'$ and when $B = B' \cup \{u\}$. Thus, we have two constraints on $\sum_{i \in B'} R_i$:

$$\sum_{i \in B'} R_i \geq H(X_{B'} | X_{\mathcal{M} \setminus B'}),$$

obtained when $B = B'$, and

$$\sum_{i \in B'} R_i \geq H(X_{B'} | X_{T \setminus B'}),$$

obtained when $B = B' \cup \{u\}$. The latter constraint is clearly stronger, so we can safely discard the former. ∎

We can now prove the desired lower bound on $R_T^{(\min)}$.

Lemma 14. *Let $m \geq 3$ be given. For $T \subset \mathcal{M}$ with $|T| = m - 1$, we have*

$$R_T^{(\min)} \geq \frac{1}{m-2} \sum_{j \in T} H(X_{T \setminus \{j\}} | X_j).$$

Proof: Consider any $T \subset \mathcal{M}$ with $|T| = m - 1$. For each $j \in T$, let $B_j = T \setminus \{j\}$. Now, let $(R_i, i \in T)$ be any point in $\mathcal{R}_T$. Applying (22) with $B = B_j$, we get

$$\sum_{i \in B_j} R_i \geq H(X_{T \setminus \{j\}} | X_j),$$

for each $j \in T$. Summing over all $j \in T$, we obtain

$$\sum_{j \in T} \sum_{i \in B_j} R_i \geq \sum_{j \in T} H(X_{T \setminus \{j\}} | X_j). \tag{23}$$

Exchanging the order of summation in the double sum on the left-hand side (LHS) above, we have

$$\sum_{j \in T} \sum_{i \in B_j} R_i = \sum_{i \in T} \sum_{j \in B_i} R_i = \sum_{i \in T} (m-2) R_i = (m-2) \sum_{i \in T} R_i.$$

Putting this back into (23), we get

$$\sum_{i \in T} R_i \geq \frac{1}{m-2} \sum_{j \in T} H(X_{T \setminus \{j\}} | X_j).$$

Since this holds for any point $(R_i, i \in T) \in \mathcal{R}_T$, the lemma follows. ∎

For the proof of Theorem 10, we need some convenient notation. For $T \subset \mathcal{M}$, $|T| = m - 1$, define $\Delta_T(\mathcal{S}) \triangleq \frac{1}{m-2}[\sum_{i \in T} H(X_i) - H(X_T)]$.

[8] Theorem 6 of [7] was based on the weaker notion of secrecy pointed out in Footnote 7. However, it can be easily verified that the result is still valid for the stronger notion of secrecy as in Definition 2.

Lemma 15. *For $m \geq 3$ terminals, if the singleton partition $\mathcal{S}$ is the unique minimizer for $\boldsymbol{I}(X_{\mathcal{M}})$, then $\Delta_T(\mathcal{S}) < \Delta(\mathcal{S})$ for all $T \subset \mathcal{M}$ with $|T| = m-1$.*

Proof: For any $u \in \mathcal{M}$, consider $T = \mathcal{M} \setminus \{u\}$. Using $\Delta(\mathcal{S}) = \frac{1}{m-1}[\sum_{i=1}^{m} H(X_i) - H(X_{\mathcal{M}})]$ and the definition of $\Delta_T(\mathcal{S})$ above, it is easy to verify the identity

$$\tfrac{m-1}{m-2}\Delta(\mathcal{S}) = \Delta_T(\mathcal{S}) + \tfrac{1}{m-2} I(X_u; X_T).$$

Re-arranging the above, we obtain

$$\begin{aligned}\Delta_T(\mathcal{S}) - \Delta(\mathcal{S}) &= \tfrac{1}{m-2}[\Delta(\mathcal{S}) - I(X_u; X_T)] \\ &= \tfrac{1}{m-2}[\Delta(\mathcal{S}) - \Delta(\mathcal{P})], \end{aligned} \tag{24}$$

where $\mathcal{P}$ is the 2-cell partition $\{\{u\}, T\}$ of $\mathcal{M}$. By assumption, the expression in (24) is strictly negative. ∎

With this, we are ready to prove Theorem 10.

Proof of Theorem 10: We will show that $\mathbf{I}(X_{\mathcal{M}}) > \mathbf{I}_T(X_{\mathcal{M}})$ for any $T \subset \mathcal{M}$ with $|T| = m-1$. First, note that since $\mathcal{S}$ is, by assumption, a minimizer for (4), we have $\mathbf{I}(X_{\mathcal{M}}) = \Delta(\mathcal{S})$. Next, by Theorem 12 and Lemma 14, we have

$$\begin{aligned}\mathbf{I}_T(X_{\mathcal{M}}) &\leq H(X_T) - \tfrac{1}{m-2}\sum_{i \in T} H(X_{T\setminus\{i\}}|X_i) \\ &= \tfrac{1}{m-2}\left[(m-2)H(X_T) - \sum_{i\in T}[H(X_T) - H(X_i)]\right] \\ &= \Delta_T(\mathcal{S}).\end{aligned}$$

Therefore, $\mathbf{I}_T(X_{\mathcal{M}}) \leq \Delta_T(\mathcal{S}) < \Delta(\mathcal{S}) = \mathbf{I}(X_{\mathcal{M}})$, the second inequality coming from Lemma 15. ∎

We conclude this section with the proof of Theorem 11. Note that when $m = 3$, (4) reduces to

$$\begin{aligned}\mathbf{I}(X_{\mathcal{M}}) = \min\{&I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2), \\ &I(X_{\{2,3\}}; X_1),\ \Delta(\mathcal{S})\},\end{aligned} \tag{25}$$

and so, the unique minimizer condition is equivalent to

$$\Delta(\mathcal{S}) < \min\{I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1)\}.$$

Note also that $\Delta(\mathcal{S}) = \frac{1}{2}[H(X_1) + H(X_2) + H(X_3) - H(X_{\{1,2,3\}})]$.

Proof of Theorem 11: The "if" part is by Theorem 10. For the "only if" part, suppose that $\Delta(\mathcal{S}) \geq \min\{I(X_{\{1,2\}}; X_3),\ I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1)\}$. Then, $\Delta(\mathcal{S})$ is either (a) greater than or equal to at least two of the three terms in the minimum, or (b) greater than or equal to exactly one term. Up to symmetry, it suffices to distinguish between two cases:

Case I: $\Delta(\mathcal{S}) \geq \max\{I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2)\}$.

Case II: $\min\{I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1)\} > \Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$.

In each case, we demonstrate a capacity-achieving communication in which at least one terminal remains silent.

We deal with Case I first. Observe that $\Delta(\mathcal{S})$ can be written as $\frac{1}{2}[I(X_1; X_2) + I(X_{\{1,2\}}; X_3)]$. Thus, the assumption $\Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$, upon some re-organization, yields $I(X_1; X_2) \geq I(X_{\{1,2\}}; X_3)$, i.e.,

$$I(X_1; X_2) \geq I(X_1; X_3) + I(X_2; X_3|X_1). \tag{26}$$

Similarly, using the identity $\Delta(\mathcal{S}) = \frac{1}{2}[I(X_1; X_3) + I(X_{\{1,3\}}; X_2)]$ in the assumption $\Delta(\mathcal{S}) \geq I(X_{\{1,3\}}; X_2)$, we obtain $I(X_1; X_3) \geq I(X_{\{1,3\}}; X_2)$, i.e.,

$$I(X_1; X_3) \geq I(X_1; X_2) + I(X_2; X_3|X_1). \tag{27}$$

The equalities in (26) and (27) can simultaneously hold iff

$$\begin{aligned} I(X_1; X_2) = I(X_1; X_3) \quad &\text{and} \\ I(X_2; X_3|X_1) = 0.& \end{aligned} \tag{28}$$

From (28), it is not hard to deduce that the quantities $I(X_{\{1,2\}}; X_3)$, $I(X_{\{1,3\}}; X_2)$ and $\Delta(\mathcal{S})$ are all equal to $I(X_1; X_2)$, and $I(X_{\{2,3\}}; X_1) = I(X_1; X_2) + I(X_1; X_3|X_2) \geq I(X_1; X_2)$. In particular, $\mathbf{I}(X_{\{1,2,3\}}) = I(X_1; X_2)$.

From the first equality in (28), we also have $H(X_1|X_2) = H(X_1|X_3)$. Now, it can be shown by a standard random binning argument that there exists a communication from terminal 1 of rate $H(X_1|X_2) = H(X_1|X_3)$ such that X_1^n is a CR. It then follows from the "balanced coloring lemma" [3, Lemma B.3] that an SK rate of $H(X_1) - H(X_1|X_2) = I(X_1; X_2)$ is achievable. Thus, the SK capacity, $\mathbf{I}(X_{\{1,2,3\}}) = I(X_1; X_2)$, is achievable by a communication in which terminals 2 and 3 are both silent.

Now, consider Case II, in which we obviously have $\mathbf{I}(X_{\{1,2,3\}}) = I(X_{\{1,2\}}; X_3)$. The idea here is to show that a valid communication of rate $H(X_{\{1,2\}}|X_3)$ exists in which terminal 3 is silent and (X_1^n, X_2^n) is a CR. Given this, an application of [3, Lemma B.3] shows that an SK rate of $H(X_{\{1,2\}}) - H(X_{\{1,2\}}|X_3) = I(X_{\{1,2\}}; X_3)$ is achievable. Thus, there is a $\mathbf{I}(X_{\{1,2,3\}})$-achieving communication in which terminal 3 is silent.

To show that the desired communication exists, we argue as follows. For $i = 1, 2$, let R_i be the rate at which terminal i communicates. A standard random binning argument shows that an achievable (R_1, R_2) region, with terminal 3 silent, for a communication intended to allow recoverability of (X_1^n, X_2^n) as CR at all terminals is given by

$$\begin{aligned} R_1 \geq H(X_1|X_2),\quad R_2 \geq H(X_2|X_1),& \\ R_1 + R_2 \geq H(X_{\{1,2\}}|X_3).& \end{aligned} \tag{29}$$

Now, using the assumption in Case II that $\Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$, we will prove that the inequality

$$H(X_1|X_2) + H(X_2|X_1) \leq H(X_{\{1,2\}}|X_3) \tag{30}$$

holds. It would then follow from (29) that there exist achievable rate pairs (R_1, R_2) with $R_1 + R_2 = H(X_{\{1,2\}}|X_3)$, thus completing the proof for Case II.

So, let us prove (30). We have $\Delta(\mathcal{S}) = \frac{1}{2}[H(X_1) + H(X_2) + H(X_3) - H(X_{\{1,2,3\}})]$ and $I(X_{\{1,2\}}; X_3) = H(X_{\{1,2\}}) + H(X_3) - H(X_{\{1,2,3\}})$. Using these expressions in the inequality $\Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$, and re-arranging terms, we obtain

$$\frac{1}{2}[H(X_1) + H(X_2) - 2H(X_{\{1,2\}})] \geq \frac{1}{2}[H(X_3) - H(X_{\{1,2,3\}})],$$

which is equivalent to (30). This completes the proof of the theorem. ∎

VI. Finding the minimizing partition

The condition that the singleton partition be a unique minimizer for $\mathbf{I}(X_{\mathcal{M}})$ plays a key role in the results of Section IV and V. Thus, it would be very useful to have a way of checking whether this condition holds for a given source $X_{\mathcal{M}}$, $m \geq 3$. The brute force method of comparing $\Delta(\mathcal{S})$ with $\Delta(\mathcal{P})$ for all partitions $\mathcal{P}$ with at least two parts requires an enormous amount of computation. Indeed, the number of partitions of an m-element set is the mth Bell number, B_m, an asymptotic estimate for which is $(\log w)^{1/2} w^{m-w} e^w$, where $w = \frac{m}{\log m}[1 + o(1)]$ is the solution to the equation $m = w \log(w+1)$ [31, Example 5.4]. The proposition below brings down the number of comparisons required for verifying the unique minimizer condition to a "mere" $2^m - m - 2$.

For any non-empty subset $B = \{b_1, b_2, \ldots, b_{|B|}\}$ of $\mathcal{M}$ with $|B| < m$, define $\mathcal{P}_B \triangleq \{B^c, \{b_1\}, \{b_2\}, \ldots, \{b_{|B|}\}\}$ to be the partition of $\mathcal{M}$ containing $|B|+1$ cells, of which $|B|$ cells are singletons comprising the elements of B. Note that if $|B| = m-1$, then $\mathcal{P}_B = \mathcal{S}$.

Proposition 16. *For $m \geq 3$, let $\Omega = \{B \subset \mathcal{M} : 1 \leq |B| \leq m-2\}$. The singleton partition $\mathcal{S}$ is*
(a) *a minimizer for $\mathbf{I}(X_{\mathcal{M}})$ iff $\Delta(\mathcal{S}) \leq \Delta(\mathcal{P}_B)\ \forall\, B \in \Omega$;*
(b) *the unique minimizer for $\mathbf{I}(X_{\mathcal{M}})$ iff $\Delta(\mathcal{S}) < \Delta(\mathcal{P}_B)\ \forall\, B \in \Omega$.*

There is in fact a strongly polynomial-time algorithm (see [12]) for determining the minimizing partition of (4). However, Proposition 16 is better suited to the purposes of our work.

Proof of Proposition 16: We prove (b); for (a), we simply have to replace the '$>$' in (31) below with a '$\geq$'.

The "only if" part is obvious. For the "if" part, suppose that $\Delta(\mathcal{S}) < \Delta(\mathcal{P}_B)$ for all $B \subset \mathcal{M}$ with $1 \leq |B| \leq m-2$. Consider any partition $\mathcal{P}$ of $\mathcal{M}$, $\mathcal{P} \neq \mathcal{S}$, with $|\mathcal{P}| \geq 2$. We wish to show that $\Delta(\mathcal{P}) > \Delta(\mathcal{S})$.

The following identity can be obtained from the definition of $\Delta(\mathcal{P})$ by some re-grouping of terms:

$$\sum_{A \in \mathcal{P}} |A^c|\, \Delta(\mathcal{P}_{A^c}) = (|\mathcal{P}|-1)[\Delta(\mathcal{P}) + (m-1)\Delta(\mathcal{S})].$$

Thus, we have

$$\begin{aligned}
\Delta(\mathcal{P}) &= \frac{1}{|\mathcal{P}|-1} \sum_{A \in \mathcal{P}} |A^c|\, \Delta(\mathcal{P}_{A^c}) - (m-1)\Delta(\mathcal{S}) \\
&> \frac{1}{|\mathcal{P}|-1} \sum_{A \in \mathcal{P}} |A^c|\, \Delta(\mathcal{S}) - (m-1)\Delta(\mathcal{S}) \qquad (31) \\
&= m\Delta(\mathcal{S}) - (m-1)\Delta(\mathcal{S}) \;=\; \Delta(\mathcal{S}). \qquad (32)
\end{aligned}$$

The inequality in (31) is due to the fact that at least one $A \in \mathcal{P}$ is not a singleton cell, so that $\mathcal{P}_{A^c} \neq \mathcal{S}$, and hence, $\Delta(\mathcal{P}_{A^c}) > \Delta(\mathcal{S})$ by assumption. To verify the first equality in (32), observe that $\sum_{A \in \mathcal{P}} |A^c| = \sum_{A \in \mathcal{P}} \sum_{i \notin A} 1 = \sum_{i=1}^m \sum_{A \in \mathcal{P}: i \notin A} 1 = m(|\mathcal{P}|-1)$. ∎

Next, we apply Proposition 16 to some interesting special cases. Random variables $X_1, X_2, \ldots, X_m$, $m \geq 2$, are called *isentropic* if $H(X_A) = H(X_B)$ for any pair of non-empty subsets $A, B \subseteq \mathcal{M}$ having the same cardinality. Equivalently, $X_1, \ldots, X_m$ are isentropic if, for all non-empty $A \subseteq \mathcal{M}$, the entropy $H(X_A)$ depends only on $|A|$. As a result, for disjoint subsets $A, B \subseteq \mathcal{M}$, conditional entropies of the form $H(X_A|X_B)$ depend only on $|A|$ and $|B|$.

Corollary 17. *Isentropic random variables form a Type $\mathcal{S}$ source.*

The proof involves checking that $\Delta(\mathcal{S}) \leq \Delta(\mathcal{P}_B)$ holds for all $B \in \Omega$, so that the result follows from Proposition 16(a). We defer the details to Appendix E.

There are many examples of isentropic random variables. For example, exchangeable random variables (cf. [32]) are isentropic. (Random variables $X_1, X_2, \ldots, X_m$ are *exchangeable* if for every permutation $\Pi : \mathcal{M} \to \mathcal{M}$, the distribution of $X_{\Pi(1)}, X_{\Pi(2)}, \ldots, X_{\Pi(m)}$ remains unchanged.) A more relevant example for us is the PIN model defined on the *complete t-uniform hypergraph on m vertices*, $K_{m,t}$. More precisely, the complete t-uniform hypergraph $K_{m,t} = (\mathcal{V}, \mathcal{E})$ has $\mathcal{V} = \mathcal{M}$, and exactly one copy of every t-subset (i.e., subset of cardinality t) of $\mathcal{M}$ belongs to $\mathcal{E}$. It is straightforward to check that the random variables $X_1, X_2, \ldots, X_m$ in the PIN model on $K_{m,t}$ are isentropic, and hence the source is Type $\mathcal{S}$. In fact, we will show below that this PIN model is strict Type $\mathcal{S}$, and therefore it satisfies the hypothesis of Theorem 10. For this and other results proved in the rest of this section, it will be useful to state a specialization of Proposition 16 to hypergraph PIN models.

In the case of hypergraph PIN models, for any $B \in \Omega$, $\Delta(\mathcal{P}_B)$ can be written as $\Delta(\mathcal{P}_B) = \frac{\sum_{e \in \mathcal{E}}[P_B(e)-1]}{|\mathcal{P}_B|-1}$, where $P_B(e)$ is the number of parts of the partition $\mathcal{P}_B$ intersecting with e. On the other hand, $\Delta(\mathcal{S}) = \frac{(t-1)|\mathcal{E}|}{m-1}$. Hence, Proposition 16 can be rewritten for the PIN model as

Corollary 18. *For a PIN model described on a t-uniform hypergraph, the singleton partition $\mathcal{S}$ is*
(a) *a minimizer for $\mathbf{I}(X_{\mathcal{M}})$ iff $\frac{(t-1)|\mathcal{E}|}{m-1} \leq \frac{\sum_{e \in \mathcal{E}}[P_B(e)-1]}{|\mathcal{P}_B|-1}$ $\forall\, B \in \Omega$;*
(b) *the unique minimizer for $\mathbf{I}(X_{\mathcal{M}})$ iff $\frac{(t-1)|\mathcal{E}|}{m-1} < \frac{\sum_{e \in \mathcal{E}}[P_B(e)-1]}{|\mathcal{P}_B|-1}$ $\forall\, B \in \Omega$.*

Corollary 19. *The PIN model on $K_{m,t}$ is strict Type $\mathcal{S}$.*

The proof is a relatively straightforward matter of checking that the condition in Corollary 18(b) holds — see Appendix E for the details.

We next give an example of a non-isentropic source which is strict Type $\mathcal{S}$. Consider the PIN model defined on a k-regular k-edge-connected graph ($t = 2$). Formally, a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called *k-regular* if every vertex in $v \in \mathcal{V}$ has degree k, i.e., there are exactly k edges in $\mathcal{E}$ which are incident with the vertex v. A graph is called *k-edge-connected* if deletion of any k-subset of $\mathcal{E}$ does not disconnect the graph, but there exists at least one $(k+1)$-subset of $\mathcal{E}$ the removal of which disconnects the graph.

Corollary 20. *A PIN model on any k-regular, k-edge-connected graph is strict Type $\mathcal{S}$.*

The proof is again an application of Corollary 18(b); the details are in Appendix E.

The m-cycle $\mathcal{C}_m$ is a special case of a k-regular and k-edge-connected graph, with $k = 2$. Formally, the m-cycle $\mathcal{C}_m = (\mathcal{V}, \mathcal{E})$, is a graph with $\mathcal{V} = \mathcal{M}$ and $\mathcal{E} = \left(\bigcup_{i=1}^{m-1}\{\{i, i+1\}\}\right) \bigcup \{\{1, m\}\}$. The complete graph $K_{m,2}$ is another example of a k-regular k-edge-connected graph with $k = m-1$. There is in fact a broad class of k-regular k-edge-connected graphs called the Harary graphs (see [33] and [34]) of which $\mathcal{C}_m$ and $K_{m,2}$ are special cases.

So far, the only example we have seen of a strict Type $\mathcal{S}$ source on a t-uniform hypergraph, with $t > 2$, has been the PIN model on the complete t-uniform hypergraph, $K_{m,t}$. It is natural to ask whether other classes of PIN models on t-uniform hypergraphs ($t > 2$) exist which are strict Type $\mathcal{S}$. The answer is 'yes'. We will construct a class of uniform hypergraphs with $t = 3$, such that the PIN models on them are strict Type $\mathcal{S}$. To do this, we introduce the *Steiner triple system* (STS) defined on the set $\mathcal{M}$. An STS on $\mathcal{M}$ is a collection of 3-subsets of $\mathcal{M}$, which we will denote by STS($\mathcal{M}$), such that any pair of elements from $\mathcal{M}$ is a subset of exactly one element of STS($\mathcal{M}$). A trivial example of an STS is $m = 3$ and STS($\mathcal{M}$) $= \{\{1, 2, 3\}\}$. It is a fact that such collections indeed exist as long as $\gcd(m-2, 6) = 1$ (see [35, Theorem 2.10]). For example, consider $m = 7$. Then, STS($\mathcal{M}$) $= \Big\{\{1,2,4\},\{2,3,5\},\{3,4,6\},\{1,5,6\},\{2,6,7\},\{1,3,7\},$ $\{4,5,7\}\Big\}$. Now, consider the 3-uniform hypergraphs $\mathcal{H}_{\text{STS}} = (\mathcal{M}, \text{STS}(\mathcal{M}))$, for all m such that STS($\mathcal{M}$) exists. We will show that a PIN model defined on $\mathcal{H}_{\text{STS}}$ with $m > 3$ is strict Type $\mathcal{S}$.

Corollary 21. *A PIN model on $\mathcal{H}_{STS}$ with $m > 3$ is strict Type $\mathcal{S}$.*

Again, the proof of the corollary is given in Appendix E.

Corollaries 19, 20 and 21 show that the PIN models on $K_{m,t}$, k-regular k-edge-connected graphs, and $\mathcal{H}_{\text{STS}}$ satisfy the hypotheses of both Theorems 6 and 10. Thus, for these sources to achieve SK capacity, an omnivocal communication is required. Also, the minimum rate of communication required is R_{CO}. Hence, in terms of public communication, these are the worst-case sources.

VII. Local Randomization

The results we have presented so far in this paper assume that the communication and the common randomness obtained from it are deterministic functions of the source. Now, Theorem 3 of [3] shows that the SK capacity is not increased even if we allow local randomization (private randomness) at the terminals. In this section, we briefly discuss how local randomization at the terminals affects the rate of communication needed to achieve SK capacity.

Consider mutually independent random variables $U_1, U_2, \ldots, U_m$, which are independent of $X_{\mathcal{M}}^n$ as well, with each terminal $i \in \mathcal{M}$ having access to U_i. As usual, U_A, $A \subseteq \mathcal{M}$ denotes the collection of random variables $(U_i : i \in A)$. We redefine the quantities in Definitions 1–6 by replacing any X_A^n, $A \subseteq \mathcal{M}$, in the definitions with the corresponding (X_A^n, U_A). We decorate the redefined quantities with a $\tilde{}$ to distinguish them from the previously defined quantities; thus, $\mathbf{I}(X_{\mathcal{M}})$ becomes $\tilde{\mathbf{I}}(X_{\mathcal{M}}, U_{\mathcal{M}})$, R_{SK} becomes $\tilde{R}_{\text{SK}}$, and so on. From Theorem 3 of [3], we have $\tilde{\mathbf{I}}(X_{\mathcal{M}}, U_{\mathcal{M}}) = \mathbf{I}(X_{\mathcal{M}})$. Moreover, since anything achievable without local randomization is also achievable with it, we have the following inequalities: $\tilde{R}_{\text{SK}} \leq R_{\text{SK}}$, $\tilde{R}_{\text{CI}} \leq R_{\text{CI}}$ and $\tilde{\text{CI}}(X_{\mathcal{M}}, U_{\mathcal{M}}) \leq \text{CI}(X_{\mathcal{M}})$. Furthermore, the proof of Theorem 2, with the necessary minor modifications, yields the inequalities $\tilde{R}_{\text{SK}} \geq \tilde{R}_{\text{CI}} \geq \tilde{\text{CI}}(X_{\mathcal{M}}, U_{\mathcal{M}}) - \mathbf{I}(X_{\mathcal{M}})$.

In the case of two terminals, Tyagi showed that local randomization has no effect on R_{SK} [8, Lemma11]. The proof first argues that $\tilde{R}_{\text{CI}} = R_{\text{CI}}$, and then leverages the fact that $R_{\text{CI}} = R_{\text{SK}}$ (in the two-terminal case) to complete the proof: $\tilde{R}_{\text{SK}} \geq \tilde{R}_{\text{CI}} = R_{\text{CI}} = R_{\text{SK}}$. The first part of Tyagi's argument can be easily extended to the multiterminal scenario, as we do next. However, we run into trouble in the second part of the argument, as we only have the inequality $R_{\text{CI}} \leq R_{\text{SK}}$ in the general multiterminal case (Theorem 2).

Proposition 22. *For a multiterminal source $X_{\mathcal{M}}^n$ with local randomization $U_{\mathcal{M}}$, we have $\tilde{R}_{CI} = R_{CI}$.*

Proof: Since $R_{\text{CI}} \geq \tilde{R}_{\text{CI}}$ holds trivially, we need to show $R_{\text{CI}} \leq \tilde{R}_{\text{CI}}$ to complete the proof. To proceed, let $(\mathbf{J}, \mathbf{F})$ be any $\tilde{\text{CI}}$. We will show that one can always construct a CI from $\mathbf{J}, \mathbf{F}$ with a lower rate of communication, and hence, $R_{\text{CI}} \leq \tilde{R}_{\text{CI}}$. Fix an $\epsilon > 0$ and let $|\mathcal{P}^*| = \ell$. It follows from (the redefined) Definition 4 that for all sufficiently large n, we have

$$\begin{aligned}
\epsilon \; &> \frac{1}{n(\ell-1)}\left[\sum_{A\in\mathcal{P}^*} H(X_A^n, U_A|\mathbf{J},\mathbf{F}) - H(X_{\mathcal{M}}^n, U_{\mathcal{M}}|\mathbf{J},\mathbf{F})\right] \\
&= \frac{1}{n(\ell-1)}\left[\sum_{A\in\mathcal{P}^*} H(U_A|\mathbf{J},\mathbf{F}) - H(U_{\mathcal{M}}|\mathbf{J},\mathbf{F})\right] \\
&\quad + \frac{1}{n(\ell-1)}\left[\sum_{A\in\mathcal{P}^*} H(X_A^n|U_A,\mathbf{J},\mathbf{F}) - H(X_{\mathcal{M}}^n|U_{\mathcal{M}},\mathbf{J},\mathbf{F})\right] \\
&\geq \frac{1}{n(\ell-1)}\left[\sum_{A\in\mathcal{P}^*} H(X_A^n|U_A,\mathbf{J},\mathbf{F}) - H(X_{\mathcal{M}}^n|U_{\mathcal{M}},\mathbf{J},\mathbf{F})\right].
\end{aligned} \tag{33}$$

The inequality in (33) is due to the fact that the term $\left[\sum_{A\in\mathcal{P}^*} H(U_A|\mathbf{J},\mathbf{F}) - H(U_{\mathcal{M}}|\mathbf{J},\mathbf{F})\right]$ is non-negative, since it can be written as the conditional relative entropy between the joint distribution of $U_{\mathcal{M}}$ and the product of the marginal distributions of $(U_A : A \in \mathcal{P}^*)$ conditioned on $\mathbf{J}, \mathbf{F}$.

Therefore, there exists at least one realization $U_{\mathcal{M}} = u_{\mathcal{M}}$ satisfying

$$\begin{aligned}
&\frac{1}{n(\ell-1)}\Bigg[\sum_{A\in\mathcal{P}^*} H(X_A^n|\mathbf{J},\mathbf{F},U_A = u_A) \\
&\qquad\qquad - H(X_{\mathcal{M}}^n|\mathbf{J},\mathbf{F},U_{\mathcal{M}} = u_{\mathcal{M}})\Bigg] \\
&= \frac{1}{n}\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|U_{\mathcal{M}} = u_{\mathcal{M}},\mathbf{J},\mathbf{F})
\end{aligned}$$

$$< \epsilon$$

for all sufficiently large n. Fixing $U_{\mathcal{M}} = u_{\mathcal{M}}$ the terminals can still agree upon $(\mathbf{J}, \mathbf{F})$, which is a CI since $\frac{1}{n}\mathbf{I}_{\mathcal{P}^*}(X_{\mathcal{M}}^n|U_{\mathcal{M}} = u_{\mathcal{M}}, \mathbf{J}, \mathbf{F}) < \epsilon$, for all sufficiently large n. Also, the range of $\mathbf{F}$ can only be lowered when conditioned on $U_{\mathcal{M}} = u_{\mathcal{M}}$, hence reducing the communication rate. Therefore, we have $R_{\text{CI}} \leq \tilde{R}_{\text{CI}}$. ∎

Since it is not clear whether the equality $R_{\text{SK}} = R_{\text{CI}}$ holds in general for a multiterminal model (see the discussion following the proof of Theorem 2 in Section III), we are unable to extend Proposition 22 to obtain the equality $\tilde{R}_{\text{SK}} = R_{\text{SK}}$. Of course, in cases where we do have the equality $R_{\text{SK}} = R_{\text{CI}}$ being true, such as in Theorem 6, we can conclude that local randomization does not affect R_{SK}.

Corollary 23. *For a Type $\mathcal{S}$ PIN model defined on a t-uniform hypergraph, we have $\tilde{R}_{SK} = R_{SK}$.*

We conjecture that the equality $\tilde{R}_{\text{SK}} = R_{\text{SK}}$ in fact holds for the general multiterminal source model, at least under some mild conditions on the amount of private randomness allowed. We are only able to prove this for "weak" SKs, which (as noted previously in Footnote 3) are defined as in our Definition 2, except that the security condition $I(\mathbf{K}; \mathbf{F}) \leq \epsilon$ is weakened to $\frac{1}{n}I(\mathbf{K}; \mathbf{F}) \leq \epsilon$. It is known [3] that the SK capacity remains unchanged even under the weaker security requirement.

Proposition 24. *For a multiterminal source $X_{\mathcal{M}}^n$ with local randomization $U_{\mathcal{M}}$ satisfying $\frac{1}{n}H(U_{\mathcal{M}}) \to 0$ as $n \to \infty$, we have for the problem of weak SK generation*

$$R_{SK} = \tilde{R}_{SK}.$$

Proof: Since, $R_{\text{SK}} \geq \tilde{R}_{\text{SK}}$ is valid trivially, we need to show that $R_{\text{SK}} \leq \tilde{R}_{\text{SK}}$. We follow an approach similar to the proof of Proposition 22. Let $\mathbf{K}$ be any maximal-rate weak SK achievable with communication $\mathbf{F}$ and local randomization $U_{\mathcal{M}}$. We will extract from $(\mathbf{K}, \mathbf{F})$ a maximal-rate weak SK and a communication achieving it without the assistance of the local randomization. The extracted communication will have a lower rate, and hence, we will have $R_{\text{SK}} = \tilde{R}_{\text{SK}}$. To proceed, fix $\epsilon > \delta > 0$. $\mathbf{K}$ being a maximal-rate SK, $\mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}) < \frac{\epsilon-\delta}{2}$ and $\frac{1}{n}I(\mathbf{K}; \mathbf{F}) < \frac{\epsilon-\delta}{2}$ hold for all sufficiently large n. Therefore, for all sufficiently large n, we have

$$\begin{aligned}
\epsilon-\delta &> \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}) + \frac{1}{n}I(\mathbf{K}; \mathbf{F}) \\
&= \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}|U_{\mathcal{M}}) + \frac{1}{n}I(\mathbf{K}; \mathbf{F}|U_{\mathcal{M}}) \\
&\quad - \frac{1}{n}I(\mathbf{K}; U_{\mathcal{M}}) + \frac{1}{n}[I(\mathbf{K}; \mathbf{F}) - I(\mathbf{K}; \mathbf{F}|U_{\mathcal{M}})] \\
&\geq \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}|U_{\mathcal{M}}) + \frac{1}{n}I(\mathbf{K}; \mathbf{F}|U_{\mathcal{M}}) - \frac{2}{n}H(U_{\mathcal{M}}) \qquad (34) \\
&\geq \mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}|U_{\mathcal{M}}) + \frac{1}{n}I(\mathbf{K}; \mathbf{F}|U_{\mathcal{M}}) - \delta, \qquad (35)
\end{aligned}$$

where (34) is because of the inequalities $|I(\mathbf{K}; \mathbf{F}) - I(\mathbf{K}; \mathbf{F}|U_{\mathcal{M}})| \leq H(U_{\mathcal{M}})$ and $I(\mathbf{K}; U_{\mathcal{M}}) \leq H(U_{\mathcal{M}})$, and (35) follows from the fact that $\frac{1}{n}H(U_{\mathcal{M}}) \to 0$ as $n \to \infty$.

Now, observe that $\mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}|U_{\mathcal{M}})$ and $\frac{1}{n}I(\mathbf{K}; \mathbf{F})$ are non-negative quantities. Thus, (35) guarantees the existence of a realization $U_{\mathcal{M}} = u_{\mathcal{M}}$, satisfying $\mathbf{I}(X_{\mathcal{M}}) - \frac{1}{n}H(\mathbf{K}|U_{\mathcal{M}} = u_{\mathcal{M}}) < \epsilon$ and $\frac{1}{n}I(\mathbf{K}; \mathbf{F}|U_{\mathcal{M}} = u_{\mathcal{M}}) < \epsilon$, for all sufficiently large n. Therefore, fixing $U_{\mathcal{M}} = u_{\mathcal{M}}$, we can choose $\mathbf{K}$ to be a maximal-rate weak SK achievable using $\mathbf{F}$ without any local randomization. Moreover, the restriction $U_{\mathcal{M}} = u_{\mathcal{M}}$ can only lower the range of $\mathbf{F}$ and hence the communication rate. Hence, we have $R_{\text{SK}} \leq \tilde{R}_{\text{SK}}$ as required. ∎

VIII. Concluding Remarks

This paper dealt with two important aspects of the public communication required to generate maximal-rate SKs in the multiterminal source model, one being the communication complexity R_{SK}, and the other being omnivocality. By extending the arguments in [8] to the setting of multiple terminals, we derived a lower bound on R_{SK} in terms of an information-theoretic quantity called the (multiterminal) interactive common information. In the two-terminal case, it was shown in [8] that this bound is always tight. Proving such a result for the general multiterminal case remains an open problem.

The minimum rate of communication for omniscience, R_{CO}, is still the best known upper bound on R_{SK}. We proved that uniform hypergraph PIN models satisfying a certain "Type $\mathcal{S}$" condition are R_{SK}-maximal. In other words, for these PIN models, R_{SK} is equal to R_{CO}. It was also shown via counterexamples that the Type $\mathcal{S}$ condition is not sufficient to guarantee R_{SK}-maximality for an arbitrary multiterminal source model. A complete characterization of R_{SK}-maximal sources is an interesting open problem.

It should be pointed out that neither our lower bound nor the R_{CO} upper bound takes into account the fact that the public communication is allowed to be *interactive*. It is possible that incorporating this information somehow leads to better bounds on R_{SK}.

The problem of characterizing communication complexity in the multiterminal source model is the stepping stone towards two bigger problems of interest. One is to characterize the communication rate region required to achieve SK capacity. The second problem is that of determining the minimum rate of communication required to generate an SK of any arbitrary rate less than or equal to SK capacity. Both these questions appear to be difficult to answer at this point. In fact, these questions are still open for the two-terminal case. It should be pointed out that these questions have been answered for a model similar to the multiterminal source model in [18]. However, that model has severe constraints on the eavesdroppers, which makes it exactly solvable but different from our model.

On the issue of omnivocality, we proved that for all strict Type $\mathcal{S}$ sources, omnivocality is needed to achieve SK capacity. The converse of this fact, i.e., omnivocality is required only if the source is strict Type $\mathcal{S}$ turns out to be true for three terminals, but no longer holds for four or more terminals. A more general problem along these lines is, given an arbitrary multiterminal source model, what is the minimum number of

terminals that must participate in a public communication to generate a maximal-rate SK for the entire set of terminals? The answer to the "dual" of this problem, i.e., what is the maximum rate of SK that can be generated when a fixed number of terminals remain silent, is already known from the work of Gohari and Anantharam [7].

APPENDIX A
PROOF OF LEMMA 7

First we state two lemmas which we will require for the proof.

Lemma 25. *For independent random variables X,Y and W, and any other random variable Z, we have*

$$I(X;Z|W) \leq I(X;Z|W,Y).$$

Proof: This follows by expanding $I(X;Y,Z|W)$ in two different ways using the chain rule, and noting that $I(X;Y|W)=0$. ∎

Lemma 26. *For independent random variables X and Y, and any other random variable Z, we have*

$$I(X;Z)+I(Y;Z) \leq I(X,Y;Z).$$

Proof: By Lemma 25, we have $I(X;Z) \leq I(X;Z|Y)$, and hence, $I(X;Z)+I(Y;Z) \leq I(X;Z|Y)+I(Y;Z) = I(X,Y;Z)$. ∎

We first show that it is enough to prove Lemma 7 for the complete t-uniform hypergraph PIN model $K_{m,t}$ (refer to Section VI for details on $K_{m,t}$) and the corresponding source $X^n_{\mathcal{M}}$. Consider any t-uniform hypergraph $\mathcal{H}=(\mathcal{V},\mathcal{E})$ with $|\mathcal{V}|=m$ and the corresponding source $\hat{X}^n_{\mathcal{M}}$, and fix a function $\mathbf{L}$ of $\hat{X}^n_{\mathcal{M}}$. For any t-subset e of $\mathcal{V}$, define $r(e)$ to be the number of times it occurs in the multiset $\mathcal{E}$, and call $r=\max\limits_{e\subset\mathcal{V}:|e|=t} r(e)$. Now, construct a new source as follows: To the multiset $\mathcal{E}^{(n)}$ add $n(r-r(e))$ copies of each t-subset e of $\mathcal{V}$. Associate with each of these newly added subsets independent Ber(1/2) random variables, which are independent of the pre-existing Ber(1/2) random variables as well. Observe that the source thus constructed is none other than $X^{nr}_{\mathcal{M}}$. Moreover, we clearly have $\sum_{i=1}^m I(X_i^{nr};\mathbf{L}) \geq \sum_{i=1}^m I(\hat{X}_i^n;\mathbf{L})$, and hence it is enough to show that $tH(\mathbf{L}) \geq \sum_{i=1}^m I(X_i^{nr};\mathbf{L})$.

For the rest of proof, we will take $X^n_{\mathcal{M}}$ to be the source described on $K_{m,t}$. We also have $I(X^n_{\mathcal{M}};\mathbf{L})=H(\mathbf{L})$ from the fact that $\mathbf{L}$ is a function of $X^n_{\mathcal{M}}$. We now show that the PIN model on $K_{m,t}$ satisfies

$$\sum_{i=1}^m I((\xi_e^n : i\in e, e\in\mathcal{E});\mathbf{L}) \;\leq\; t\,I((\xi_e^n : e\in\mathcal{E});\mathbf{L}), \quad (36)$$

where ξ_e^n represents the collection of the n ξ_e's associated with the n copies of the hyperedge e in $\mathcal{E}^{(n)}$.

For any $i\in\mathcal{M}$, let $\mathcal{E}_i$ denote the set of hyperedges containing i, so that the left-hand side of (36) can be expressed as $\sum_{i=1}^m I((\xi_e^n : e\in\mathcal{E}_i);\mathbf{L})$. Now, we write $\mathcal{E}_i$ as a union of two disjoint sets $\mathcal{E}_{\geq i}$ and $\mathcal{E}_{\not> i}$, i.e., $\mathcal{E}_i=\mathcal{E}_{\geq i}\bigcup\mathcal{E}_{\not> i}$. The set $\mathcal{E}_{\geq i}$ is the subset of $\mathcal{E}_i$ containing no terminals from $\{1,2,\ldots,i-1\}$. The set $\mathcal{E}_{\not> i}$ is thus the subset of $\mathcal{E}_i$ containing at least one terminal from $\{1,2,\ldots,i-1\}$. Observe that we have $|\mathcal{E}_{\geq i}|=\binom{m-i}{t-1}$ for $1\leq i\leq m-t+1$ and $|\mathcal{E}_{\geq i}|=0$ for $m-t+2\leq i\leq m$. Therefore,

$$\begin{aligned}
&\sum_{i=1}^m I((\xi_e^n : e\in\mathcal{E}_i);\mathbf{L})\\
&= I\left((\xi_e^n : e\in\mathcal{E}_{\geq 1});\mathbf{L}\right) + \sum_{i=2}^{m-t+1}\Big[I\left(\left(\xi_e^n : e\in\mathcal{E}_{\not> i}\right);\mathbf{L}\right)\\
&\quad + I\left(\left(\xi_e^n : e\in\mathcal{E}_{\geq i}\right);\mathbf{L}\Big|\left(\xi_e^n : e\in\mathcal{E}_{\not> i}\right)\right)\Big]\\
&\quad + \sum_{i=m-t+2}^{m} I\left((\xi_e^n : e\in\mathcal{E}_i);\mathbf{L}\right)\\
&\leq I\left((\xi_e^n : e\in\mathcal{E}_{\geq 1});\mathbf{L}\right)\\
&\quad + \sum_{i=2}^{m-t+1} I\left(\left(\xi_e^n : e\in\mathcal{E}_{\geq i}\right);\mathbf{L}\Big|\left(\xi_e^n : e\in\bigcup_{j\leq i}\mathcal{E}_{\not> j}\right)\right)\\
&\quad + \sum_{i=2}^{m-t+1} I\left(\left(\xi_e^n : e\in\mathcal{E}_{\not> i}\right);\mathbf{L}\right)\\
&\quad + \sum_{i=m-t+2}^{m} I\left((\xi_e^n : e\in\mathcal{E}_i);\mathbf{L}\right) \qquad (37)\\
&= \underbrace{I\left(\left(\xi_e^n : e\in\mathcal{E}\right);\mathbf{L}\right)}_{P} + \underbrace{\sum_{i=2}^{m-t+1} I\left(\left(\xi_e^n : e\in\mathcal{E}_{\not> i}\right);\mathbf{L}\right)}_{Q}\\
&\quad + \underbrace{\sum_{i=m-t+2}^{m} I\left((\xi_e^n : e\in\mathcal{E}_i);\mathbf{L}\right)}_{R}, \qquad (38)
\end{aligned}$$

where (37) follows from Lemma 25. Note that for $t=2$, (36) follows directly from (38): by virtue of Lemma 26, we have $Q+R\leq P$, so that the right-hand side (RHS) of (38) is at most $2P$, as desired. However, the case of $t>2$ is not as simple and needs further work.

To achieve the RHS of (36), we require $Q+R\leq(t-1)P$. We proceed by defining $Q(i)=I\left(\left(\xi_e^n : e\in\mathcal{E}_{\not> i}\right);\mathbf{L}\right)$ for all $2\leq i\leq m-t+1$, and thus, $Q=\sum_{i=2}^{m-t+1}Q(i)$. Similarly, define $R(i)=I\left((\xi_e^n : e\in\mathcal{E}_i);\mathbf{L}\right)$ for all $m-t+2\leq i\leq m$, so that $R=\sum_{i=m-t+2}^{m}R(i)$. The key ideas are the following:

1) Expand each $Q(i)$ using the chain rule into conditional mutual information terms of the form $I(\xi_e^n;\mathbf{L}|\cdots)$, and further condition them on additional $\xi_{\tilde{e}}^n$s appropriately.
2) Allocate these conditional mutual information terms to appropriate $R(i)$s.
3) Use the chain rule to sum each $R(i)$ and the terms allocated to it to obtain P.

Since the conditional mutual information term $I(\xi_e^n;\mathbf{L}|\cdots)$ can only increase upon further conditioning on additional $\xi_{\tilde{e}}^n$s (by Lemma 25), we have $Q+R\leq(t-1)P$ as required.

To proceed, we need to define a total ordering on the set $\mathcal{E}$. We represent a hyperedge e as a t-tuple $(i_1 i_2 \ldots i_t)$, with the i_js, $1\leq j\leq t$, being the terminals which are contained in e, ordered according to $i_1<i_2<\ldots<i_t$. We will use '$<$' to denote the lexicographic ordering of the t-tuples (hyperedges)

in $\mathcal{E}$. Furthermore, based on the ordering '$<$', we index the hyperedges of $\mathcal{E}$ as e_j, $1 \leq j \leq \binom{m}{t}$, satisfying $e_i < e_j$ iff $i < j$. As an example, Table I illustrates the indexing of the hyperedges in $K_{5,3}$.

TABLE I: Indexing of the hyperedges in $K_{5,3}$

Hyperedge	Index
(123)	1
(124)	2
(125)	3
(134)	4
(135)	5
(145)	6
(234)	7
(235)	8
(245)	9
(345)	10

To proceed further, using the chain rule we expand each $Q(i)$ into a sum of conditional mutual information terms of the form $Q_e \triangleq I(\xi_e^n; \mathbf{L} | (\xi_{\tilde{e}}^n : \tilde{e} < e, \tilde{e} \in \mathcal{E}))$ as follows:

$$\begin{aligned} Q(i) &= I((\xi_e^n : e \in \mathcal{E}_{\not\ni i}); \mathbf{L}) \\ &= \sum_{e \in \mathcal{E}_{\not\ni i}} I(\xi_e^n; \mathbf{L} | (\xi_{\tilde{e}}^n : \tilde{e} < e, \tilde{e} \in \mathcal{E}_{\not\ni i})) \\ &\leq \sum_{e \in \mathcal{E}_{\not\ni i}} I(\xi_e^n; \mathbf{L} | (\xi_{\tilde{e}}^n : \tilde{e} < e, \tilde{e} \in \mathcal{E})) \qquad (39) \\ &= \sum_{e \in \mathcal{E}_{\not\ni i}} Q_e, \qquad (40) \end{aligned}$$

where (39) follows from Lemma 25. Hence, we have $Q \leq \sum_{i=2}^{m-t+1} \sum_{e \in \mathcal{E}_{\not\ni i}} Q_e$. A total of $\sum_{i=2}^{m-t+1} \left[\binom{m-1}{t-1} - \binom{m-i}{t-1} \right] = (t-1)\binom{m-1}{t} Q_e$ terms are generated. Next, each $R(i)$ is allocated $\binom{m-1}{t}$ terms Q_{e_j}, $1 \leq j \leq \binom{m}{t}$, satisfying $i \notin e_j$. This allocation procedure is explained in detail below and is also formalized in Algorithm 1. We add a further conditioning on each Q_{e_j} allocated to $R(i)$ to make it $Q_{e_{j|i}} \triangleq I(\xi_{e_j}^n; \mathbf{L} | (\xi_{\tilde{e}}^n : \tilde{e} < e_j, \tilde{e} \in \mathcal{E}), (\xi_{\tilde{e}}^n : \tilde{e} \in \mathcal{E}_i))$. Lemma 25 and the definition of $Q_{e_{j|i}}$ ensure that $R(i) + \sum_{j: i \notin e_j} Q_{e_j} \leq R(i) + \sum_{j: i \notin e_j} Q_{e_{j|i}} = P$.

We now give a more detailed description of the allocation procedure. Construct a table T with rows indexed by $i = 2, 3, \ldots, m-t+1$ and the columns indexed by $j = 1, 2, \ldots, \binom{m}{t}$. This table records the availability (for allocation) of a Q_{e_j} from the expansion of $Q(i)$ in (40). Initialize the table as follows: $T(i,j) = 1$ if a Q_{e_j} came from $Q(i)$ in (40); else $T(i,j) = 0$. We carry out the allocation procedure on each $R(i)$ in ascending order of i. The procedure of allocation is as follows. The idea is to allocate the necessary Q_{e_j}s to $R(i)$ in ascending order of j. Once an i and e_j are fixed, we test whether $i \notin e_j$ is satisfied. If not, we increment j by 1. If $i \notin e_j$ is satisfied, then the availability of Q_{e_j} from $Q(k)$, for all $2 \leq k \leq m-t+1$, is checked using the table T. The smallest k which satisfies $T(k,j) = 1$ is chosen, and $R(i)$ is allocated the Q_{e_j} coming from that $Q(k)$. The table is then updated with $T(k,j) = 0$ to record that the Q_{e_j} from that $Q(k)$ is no longer available for allocation. We then increment j by 1 and repeat the allocation procedure. Once all Q_{e_j}s with $i \notin e_j$ have been allocated to $R(i)$, we begin the allocation procedure for $R(i+1)$. We formally summarize this allocation procedure in Algorithm 1.

Algorithm 1

$i = m-t+2, j = 1$.
while $i \leq m$ **do**
  if $i \notin e_j$ **then**
    $k = 2$.
    while $k \leq m-t+1$ **do**
      if $T(k,j) = 1$ **then**
      Choose the Q_{e_j} coming from $Q(k)$ in (40).
      Add the additional conditioning to make it $Q_{e_{j|i}}$.
      Allocate this term to $R(i)$.
      $T(k,j) \leftarrow 0$.
      Break.
      end if
      if $T(k,j) = 0$ && $k = m-t+1$ **then**
        Declare ERROR and halt.
      end if
      $k \leftarrow k+1$.
    end while
  end if
  $j \leftarrow j+1$.
  if $j = \binom{m}{t} + 1$ **then**
    $i \leftarrow i+1$.
    $j \leftarrow 1$.
  end if
end while

The flow of Algorithm 1 for $K_{5,3}$ is illustrated in Example A.1 further below. We now make the following claims:

Claim 1. *Algorithm 1 never terminates in ERROR.*

Claim 2. *Algorithm 1 exhausts all the Q_e terms generated in* (40).

Claim 1 ensures that each $R(i)$, for all $m-t+2 \leq i \leq m$, is allocated all the Q_{e_j}s satisfying $i \notin e_j$. Therefore, using Claim 2, we have

$$\begin{aligned} Q + R &= \sum_{i=m-t+2}^{m} \left[R(i) + \sum_{j: i \notin e_j} Q_{e_j} \right] \\ &\leq \sum_{i=m-t+2}^{m} \left[R(i) + \sum_{j: i \notin e_j} Q_{e_{j|i}} \right] = (t-1)P. \end{aligned}$$

This completes the proof of Lemma 7, modulo the proofs of Claims 1 and 2, which we give below.

Proof of Claim 1: ERROR is possible only if for some $m-t+2 \leq i \leq m$ and for some e satisfying $i \notin e$, all the Q_e terms generated in (40) have already been allocated. This is impossible as there are always enough Q_es. To see this, suppose e contains $t-1-p$ terminals from $\{m-t+2, \ldots, m\}$, i.e., there are p $R(i)$s requiring an allocation of Q_e. Since the hypergraph is t-uniform, e must contain $p+1$ terminals from $\{1, 2, \ldots, m-t+1\}$. This implies that the total number of

Q_es generated in (40) is p. Therefore, we clearly have enough Q_es for all $R(i)$s. ∎

Proof of Claim 2: As discussed earlier, the total number of Q_e terms generated in (40) is $(t-1)\binom{m-1}{t}$. Also, the total number of Q_e terms required by each $R(i)$ is $\binom{m-1}{t}$. Therefore, using Claim 1, the claim follows. ∎

Example A.1. *We illustrate how Algorithm 1 proceeds for $K_{5,3}$. Denote the hyperedges in $\mathcal{E}$ using 3-tuples, i.e., the hyperedge containing terminals 1, 2 and 3 is (123). The indexing of $\mathcal{E}$ is illustrated in Table I. So for this case we have $Q(2) = I(\xi^n_{(123)}, \xi^n_{(124)}, \xi^n_{(125)}; \boldsymbol{L})$ and $Q(3) = I(\xi^n_{(123)}, \xi^n_{(134)}, \xi^n_{(135)}, \xi^n_{(234)}, \xi^n_{(235)}; \boldsymbol{L})$. Thus,* (40) *takes the form*

$$\begin{aligned} Q(2) \;\leq\; & I(\xi^n_{(123)}; \boldsymbol{L}) + I(\xi^n_{(124)}; \boldsymbol{L} | (\xi^n_e : e < (124)) \\ & + I(\xi^n_{(125)}; \boldsymbol{L} | (\xi^n_e : e < (125)), \qquad (41) \\ Q(3) \leq\; & I(\xi^n_{(123)}; \boldsymbol{L}) + I(\xi^n_{(134)}; \boldsymbol{L} | (\xi^n_e : e < (134)) \\ & + I(\xi^n_{(135)}; \boldsymbol{L} | (\xi^n_e : e < (135)) \\ & + I(\xi^n_{(234)}; \boldsymbol{L} | (\xi^n_e : e < (234)) \\ & + I(\xi^n_{(235)}; \boldsymbol{L} | (\xi^n_e : e < (235)). \qquad (42) \end{aligned}$$

Observe that $R(4)$ and $R(5)$ require four Q_e terms each, and a total of eight Q_e terms are in fact available from (41) *and* (42)*. The table T is initialized as follows:*

	1	2	3	4	5	6	7	8	9	10
2	1	1	1	0	0	0	0	0	0	0
3	1	0	0	1	1	0	1	1	0	0

We will now illustrate a few of the allocations carried out by Algorithm 1. The algorithm begins with $i = 4$ and $j = 1$ and $Q_{(123)}$ needs to be allocated to $R(4)$. With $k = 2$ we see that $T(k, 1) = 1$, and hence we allocate $Q_{(123)}$ coming from $Q(2)$ to $R(4)$. The table T is then updated as below.

	1	2	3	4	5	6	7	8	9	10
2	0	1	1	0	0	0	0	0	0	0
3	1	0	0	1	1	0	1	1	0	0

Next we will illustrate the allocation of $Q_{(123)}$ to $R(5)$, i.e., $i = 5$ and $j = 1$. The state of the table T just before this step is shown below.

	1	2	3	4	5	6	7	8	9	10
2	0	1	0	0	0	0	0	0	0	0
3	1	0	0	1	0	0	1	0	0	0

Setting $k = 2$, we see that $T(k, 1) = 0$. So, we move to $k = 3$, for which $T(k, 1) = 1$. Hence the $Q_{(123)}$ term coming from $Q(3)$ is allocated to $R(5)$, and the table T is updated as below.

	1	2	3	4	5	6	7	8	9	10
2	0	1	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	1	0	0	0

We give one last example of an allocation. Observe that $e = (234)$ is the largest (in terms of the ordering on $\mathcal{E}$) hyperedge such that Q_e needs to be allocated to $R(5)$. We will now illustrate this step. This happens when $i = 5$ and $j = 7$. The updated table T just before this step is shown below.

	1	2	3	4	5	6	7	8	9	10
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	1	0	0	0

With $k = 2$, we see that $T(k, 7) = 0$. So set $k = 3$, and note that $T(k, 7) = 1$. So, we allocate to $R(5)$ the $Q_{(234)}$ term contributed by $Q(3)$. Upon updating, the table T now has all entries to be 0. Observe that at this point no other allocation is required, as the Q_{e_j}s for $j = 8$, 9 and 10 are not required by $R(5)$ since terminal 5 is contained in each of e_8, e_9 and e_{10}. Thus Algorithm 1 successfully terminates. Finally, we rewrite (41) *and* (42) *with underbraces showing the $R(i)$ term to which each Q_e term was allocated by Algorithm 1.*

$$\begin{aligned} Q(2) \;\leq\; & \underbrace{I(\xi^n_{(123)}; \boldsymbol{L})}_{R(4)} + \underbrace{I(\xi^n_{(124)}; \boldsymbol{L} | (\xi^n_e : e < (124))}_{R(5)} \\ & + \underbrace{I(\xi^n_{(125)}; \boldsymbol{L} | (\xi^n_e : e < (125))}_{R(4)} \qquad (43) \\ Q(3) \;\leq\; & \underbrace{I(\xi^n_{(123)}; \boldsymbol{L})}_{R(5)} + \underbrace{I(\xi^n_{(134)}; \boldsymbol{L} | (\xi^n_e : e < (134))}_{R(5)} \\ & + \underbrace{I(\xi^n_{(135)}; \boldsymbol{L} | (\xi^n_e : e < (135))}_{R(4)} \\ & + \underbrace{I(\xi^n_{(234)}; \boldsymbol{L} | (\xi^n_e : e < (234))}_{R(5)} \\ & + \underbrace{I(\xi^n_{(235)}; \boldsymbol{L} | (\xi^n_e : e < (235))}_{R(4)} \qquad (44) \end{aligned}$$

It can be clearly seen from (43) *and* (44) *that $R(i), i = 4, 5$, have each been allocated with all Q_es with $i \notin e$, and no Q_e is left unallocated.*

Appendix B
The proof of Lemma 9

Fix an $n \in \mathbb{N}$ and let $\{T_1, T_2, \ldots, T_{\sigma^{(n)}}\}$ be a set of edge-disjoint spanning trees of $\mathcal{G}^{(n)}$ of maximum cardinality $\sigma^{(n)} := \sigma(\mathcal{G}^{(n)})$. We will run Protocol 1 of [32] independently on each of the trees $T_j, 1 \leq j \leq \sigma^{(n)}$. For the sake of completeness, we describe the protocol below.

Fix a spanning tree $T_j, 1 \leq j \leq \sigma^{(n)}$, and fix a specific edge e from the set of edges of T_j. Define $\xi(T_j) := \xi_e$, where, as usual, ξ_e denotes the random variable associated with the edge e. For any vertex $i \in \mathcal{M}$, denote by $d_j(i)$ the degree of the vertex i in the spanning tree T_j. For any vertex i satisfying $d_j(i) > 1$, without loss of generality we label the edges of T_j incident on it by $e(1), e(2), \ldots, e(d)$, where $d = d_j(i)$. The communication from terminal i derived from T_j is $\mathbf{F}_{T_j}(i) := (\xi_{e(1)} \oplus \xi_{e(2)}, \xi_{e(2)} \oplus \xi_{e(3)}, \ldots, \xi_{e(d-1)} \oplus \xi_{e(d)})$, where $\oplus$ denotes the modulo-2 sum. Let $\mathbf{F}_{T_j} = \big(\mathbf{F}_{T_j}(1), \mathbf{F}_{T_j}(2), \ldots, \mathbf{F}_{T_j}(m)\big)$, and let $\mathcal{F}_{T_j}$ denote the range of $\mathbf{F}_{T_j}$. It is not hard to check the following facts: Firstly, every terminal can recover $\xi(T_j)$ from $\mathbf{F}_{T_j}$. Secondly, $I(\mathbf{F}_{T_j}; \xi(T_j)) = 0$. Thirdly,

$$\log|\mathcal{F}_{T_j}| \;=\; \sum_{i=1}^{m} [d_j(i) - 1] \;=\; \sum_{i=1}^{m} d_j(i) - m$$

$$= \ 2(m-1) - m \ = \ m-2, \tag{45}$$

where we have used the fact that the number of edges in any spanning tree is $m-1$.

To complete the proof, we show that this protocol has communication rate $(m-2)\overline{\sigma}(\mathcal{G})$ and achieves SK capacity. Denote the entire communication $(\mathbf{F}_{T_1}, \mathbf{F}_{T_2}, \ldots, \mathbf{F}_{T_{\sigma^{(n)}}})$ by $\mathbf{F}$ and denote its range by $\mathcal{F}$. Set $\mathbf{K} = (\xi(T_1), \xi(T_2), \ldots, \xi(T_{\sigma^{(n)}}))$. Noting that the spanning trees $T_j, 1 \leq j \leq \sigma^{(n)}$ are edge-disjoint, we have, using the independence of the random variables associated with the edges in $\mathcal{E}^{(n)}$, $H(\mathbf{K}) = \sigma^{(n)}$, $\log|\mathcal{F}| = (m-2)\sigma^{(n)}$ and $I(\mathbf{K};\mathbf{F}) = 0$. Therefore, $\mathbf{K}$ is a secret key satisfying $\lim\limits_{n\to\infty} \frac{1}{n} H(\mathbf{K}) = \overline{\sigma}(\mathcal{G})$, and hence the protocol is capacity-achieving. The protocol used a communication rate of $(m-2)\overline{\sigma}(\mathcal{G})$ and thus $R_{\text{SK}} \leq (m-2)\overline{\sigma}(\mathcal{G})$.

Appendix C
An example of a non-R_{SK}-maximal strict Type $\mathcal{S}$ source

In this section we provide an example of a source which is strict Type $\mathcal{S}$ and yet is non R_{SK}-maximal. To construct such a source we need to define "clubbing together" of independent multiterminal sources on $\mathcal{M}$. Formally, for independent sources $X^n_{\mathcal{M}}$ and $Y^n_{\mathcal{M}}$, define the *clubbed* source $Z^n_{\mathcal{M}}$ as $Z^n_i = (X^n_i, Y^n_i)$, for all $i \in \mathcal{M}$. Π^*_X and Π^*_Y are defined to be the sets of partitions of $\mathcal{M}$ which are minimizers of (4) for $X^n_{\mathcal{M}}$ and $Y^n_{\mathcal{M}}$, respectively. We will denote the communication complexity (resp. minimum rate of communication for omniscience) for the individual sources $X^n_{\mathcal{M}}$ and $Y^n_{\mathcal{M}}$ by R_{SK_X} and R_{SK_Y} (resp. R_{CO_X} and R_{CO_Y}) respectively. The clubbed source satisfies the following result.

Proposition 27. *Consider two independent multiterminal sources $X^n_{\mathcal{M}}$ and $Y^n_{\mathcal{M}}$ and the corresponding clubbed source $Z^n_{\mathcal{M}}$. Then we have*

$$\boldsymbol{I}(Z_{\mathcal{M}}) \ \geq \ \boldsymbol{I}(X_{\mathcal{M}}) + \boldsymbol{I}(Y_{\mathcal{M}}) \tag{46}$$

with equality iff $\Pi^*_X \bigcap \Pi^*_Y \neq \emptyset$.

Proof: Consider any partition $\mathcal{P} = \{A_1, A_2, \cdots, A_\ell\}$ of $\mathcal{M}$. We have

$$\begin{aligned}
\Delta(\mathcal{P}) \ &= \ \frac{1}{\ell-1}\left[\sum_{i=1}^{\ell} H(Z_{A_i}) - H(Z_{\mathcal{M}})\right] \\
&= \ \underbrace{\frac{1}{\ell-1}\left[\sum_{i=1}^{\ell} H(X_{A_i}) - H(X_{\mathcal{M}})\right]}_{\Delta_X(\mathcal{P})} \\
&\quad + \underbrace{\frac{1}{\ell-1}\left[\sum_{i=1}^{\ell} H(Y_{A_i}) - H(Y_{\mathcal{M}})\right]}_{\Delta_Y(\mathcal{P})}, \qquad (47)
\end{aligned}$$

where (47) follows from the independence of $X^n_{\mathcal{M}}$ and $Y^n_{\mathcal{M}}$.

Thus we have from (47) that $\min_{\mathcal{P}} \Delta(\mathcal{P}) \geq \min_{\mathcal{P}} \Delta_X(\mathcal{P}) + \min_{\mathcal{P}} \Delta_Y(\mathcal{P})$ with equality iff $\mathcal{P} \in \Pi^*_X \bigcap \Pi^*_Y$. The result follows. ∎

We conclude the section by constructing a non R_{SK}-maximal source with $\mathcal{S}$ being the unique minimizer in (4).

Example C.1. *Consider a clubbed source $Z^n_{\mathcal{M}} = (X^n_{\mathcal{M}}, Y^n_{\mathcal{M}})$, where $X^n_{\mathcal{M}}$ is the source described in Example IV.1 and $Y^n_{\mathcal{M}}$ corresponds to the PIN model on a k-regular, k-edge-connected graph. By Corollary 20, we have $\Pi^*_Y = \{\mathcal{S}\}$.*

Since $\Pi^*_X \bigcap \Pi^*_Y = \{\mathcal{S}\}$, using Proposition 27 we have SK capacity $\boldsymbol{I}(Z_{\mathcal{M}}) = \boldsymbol{I}(X_{\mathcal{M}}) + \boldsymbol{I}(Y_{\mathcal{M}})$. By independently running protocols achieving R_{SK_X} and R_{SK_Y}, an SK of rate $\boldsymbol{I}(X_{\mathcal{M}}) + \boldsymbol{I}(Y_{\mathcal{M}})$, i.e., SK capacity can be achieved. The communication rate used in independently running the two protocols is $R_{SK_X} + R_{SK_Y}$. Now, (2) and the independence of $X^n_{\mathcal{M}}$ and $Y^n_{\mathcal{M}}$ show that $R_{CO} = R_{CO_X} + R_{CO_Y}$. On the other hand, it is shown in Example IV.1 that $R_{SK_X} < R_{CO_X}$. Therefore, we have

$$R_{SK} \ \leq \ R_{SK_X} + R_{SK_Y} \ < \ R_{CO_X} + R_{CO_Y} \ = \ R_{CO}.$$

Appendix D
A non-strict Type $\mathcal{S}$ source requiring omnivocality[9]

For $m \geq 4$, consider the multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \mathcal{M}$ as usual. The multiset $\mathcal{E}$ consists of $m-2$ copies of the edges $\{i, i+1\}$ for $1 \leq i \leq m-1$, and $m-1$ copies of the edge $\{1, m\}$. Using techniques derived in [6, Theorem 5], it can be shown that for the PIN model defined on $\mathcal{G}$, we have $\mathbf{I}(X_{\mathcal{M}}) = m-1$. We will show below that this PIN model is non-strict Type $\mathcal{S}$, and yet it requires omnivocality to achieve SK capacity.

We first show that the source is not strict Type $\mathcal{S}$. Simple computations reveal the following facts: $H(X_i) = 2(m-2)$, for all $i \in \{2, 3, \ldots, m-1\}$, $H(X_1) = H(X_m) = 2(m-2)+1$, $H(X_1, X_m) = 3(m-2)+1$ and $H(X_{\mathcal{M}}) = m(m-2)+1$. Using these it is easy to check that $\Delta(\mathcal{S}) = m-1$, and moreover, $\Delta(\mathcal{P}') = m-1$, where $\mathcal{P}' = \{\{1, m\}, \{2\}, \{3\}, \ldots, \{m-1\}\}$. Hence the source $X^n_{\mathcal{M}}$ is Type $\mathcal{S}$, but not strict Type $\mathcal{S}$.

Now, we show that this source requires omnivocality to achieve SK capacity. As in the proof of Theorem 10, we make use of Theorem 12, and show that for any $T \subset \mathcal{M}$ with $|T| = m-1$, we have $\mathbf{I}_T(X_{\mathcal{M}}) < \mathbf{I}(X_{\mathcal{M}})$. Let $T = \mathcal{M} \setminus \{u\}$ with $u \in \mathcal{M}$. Using symmetry it is enough to show $\mathbf{I}_T(X_{\mathcal{M}}) < \mathbf{I}(X_{\mathcal{M}})$ for the following two cases:

Case I: $u = 1$.

Case II: $u \in \{2, 3, \ldots, m-2\}$.

In both cases we will derive lower bounds on $R_T^{(\min)}$ and hence obtain an upper bound on $\mathbf{I}_T(X_{\mathcal{M}})$. First we deal with Case I with $T = \{2, 3, \ldots, m\}$. In this case, $H(X_T) = m(m-2)+1$. Also, any point in $\mathcal{R}_T$ satisfies the following constraints from (22):

$$\begin{aligned}
\sum_{i=2}^{m-1} R_i \ &\geq \ H(X_2, X_3, \ldots, X_{m-1} | X_m) \ = \ (m-2)^2, \\
R_m \ &\geq \ H(X_m | X_2, X_3, \ldots, X_{m-1}) \ = \ m-1.
\end{aligned}$$

[9]This example is a contribution of Chan et al. See [12, extension to Example C.1].

Using the above constraints, we have $R_T^{(\min)} \geq (m-1) + (m-2)^2$. Thus,

$$\begin{aligned}\mathbf{I}_T(X_{\mathcal{M}}) &= H(X_T) - R_T^{(\min)} \\ &\leq m(m-2)+1-(m-1)-(m-2)^2 \\ &= m-2 \\ &< m-1 = \mathbf{I}(X_{\mathcal{M}}). \end{aligned} \tag{48}$$

Hence, SK capacity cannot be achieved with terminal 1 remaining silent.

Next we deal with Case II. Assume an arbitrary $u \in \{2,3,\ldots,m-2\}$ is silent. As in Case I, we have $H(X_T) = m(m-2)+1$. We see from (22) that the rate region $\mathcal{R}_T$ is defined in part by the following constraints:

$$\begin{aligned}&\sum_{i=1}^{u-1} R_i + R_m \\ &\quad\geq H(X_1, X_2, \ldots, X_{u-1}, X_m | X_{u+1}, X_{u+2}, \ldots, X_{m-1}) \\ &\quad= u(m-2)+1,\end{aligned}$$

$$\begin{aligned}&\sum_{i=u+1}^{m-1} R_i \\ &\quad\geq H(X_{u+1}, X_{u+2}, \ldots, X_{m-1} | X_1, X_2, \ldots, X_{u-1}, X_m) \\ &\quad= (m-u-1)(m-2).\end{aligned}$$

The above constraints imply that $R_T^{(\min)} \geq (m-2)(m-1)+1 = (m-2)^2 + (m-1)$. Hence, as in (48), we have $\mathbf{I}_T(X_{\mathcal{M}}) < \mathbf{I}(X_{\mathcal{M}})$.

Therefore, the source $X_{\mathcal{M}}^n$ cannot attain SK capacity without using omnivocality.

Appendix E
Proofs of Corollaries of Proposition 16

In this section, we give the proofs of Corollaries 17, 19, 20 and 21. We start with the corollary stating that isentropic random variables form a Type $\mathcal{S}$ source.

Proof of Corollary 17: For a partition $\mathcal{P}$ of $\mathcal{M}$ with $|\mathcal{P}| \geq 2$, let us define

$$\delta(\mathcal{P}) \triangleq \frac{1}{|\mathcal{P}|-1} \sum_{A \in \mathcal{P}} H(X_{A^c}|X_A) = H(X_{\mathcal{M}}) - \Delta(\mathcal{P}).$$

By virtue of Proposition 16(a), we need to show that $\delta(\mathcal{P}_B) \leq \delta(\mathcal{S})$ for all $B \in \Omega$.

For isentropic random variables, the quantity $H(X_B|X_{B^c})$, for any $B \subseteq \mathcal{M}$, depends only on the cardinality of B. Thus, for $1 \leq k \leq m$, define $g(k) \triangleq H(X_{\{1,2,\ldots,k\}}|X_{\mathcal{M}\setminus\{1,2,\ldots,k\}})$; also, set $g(0)=0$. With this, we can write

$$\begin{aligned}\delta(\mathcal{P}_B) &= \frac{1}{|B|}\left[H(X_B|X_{B^c}) + \sum_{i\in B} H(X_{\mathcal{M}\setminus\{i\}}|X_i)\right] \\ &= \frac{1}{|B|} g(|B|) + g(m-1).\end{aligned}$$

Also, note that $\delta(\mathcal{S}) = \frac{m}{m-1} g(m-1)$. Thus, we have to show that $\frac{g(|B|)}{|B|} \leq \frac{g(m-1)}{m-1}$ for all $B \in \Omega$. We accomplish this by proving that for isentropic random variables, the function $g(k)/k$ is non-decreasing in k, or equivalently, $kg(k+1) - (k+1)g(k)$ is always non-negative. Indeed, we have $g(k+1) = H(X_{\mathcal{M}}) - H(X_{\{k+2,\ldots,m\}})$ and $g(k) = H(X_{\mathcal{M}}) - H(X_{\{k+1,\ldots,m\}}) = g(k+1) - H(X_{k+1}|X_{\{k+2,\ldots,m\}})$. Thus,

$$\begin{aligned}&kg(k+1) - (k+1)g(k) \\ &\qquad= (k+1)\, H(X_{k+1}|X_{\{k+2,\ldots,m\}}) - g(k+1).\end{aligned}$$

It is straightforward to show that the above quantity is non-negative:

$$\begin{aligned}g(k+1) &= H(X_{\{1,2,\ldots,k+1\}}|X_{\{k+2,\ldots,m\}}) \\ &\leq \sum_{i=1}^{k+1} H(X_i|X_{\{k+2,\ldots,m\}}) \\ &= (k+1)H(X_{k+1}|X_{\{k+2,\ldots,m\}}),\end{aligned}$$

since, for $1 \leq i \leq k+1$, $H(X_i|X_{\{k+2,\ldots,m\}}) = H(X_{k+1}|X_{\{k+2,\ldots,m\}})$ by isentropy. ∎

Next, we prove Corollary 19, which states that the PIN model on $K_{m,t}$ is strict Type $\mathcal{S}$.

Proof of Corollary 19: Fix a set $B \subsetneq \mathcal{M}$ with $|B| \leq m-2$. We will use Corollary 18 to show that the PIN model on $K_{m,t}$ is strict Type $\mathcal{S}$. First we make the observation that $|\mathcal{E}| = \binom{m}{t}$ for the case of $K_{m,t}$. To proceed, we need to evaluate the expression $\sum_{e\in\mathcal{E}}[P_B(e)-1]$. We first consider the case when $|B| \geq t$. The fact that $|B|$ is at least t implies that there are $\binom{|B|}{t}$ hyperedges which contain only elements of B, i.e., intersect the partition $\mathcal{P}_B$ in t parts. Now fix an i with $1 \leq i \leq t-1$. There are $\binom{|B|}{i}\binom{m-|B|}{t-i}$ hyperedges containing any i terminals from B and any $t-i$ terminals from $\mathcal{M}\setminus B$, i.e., intersecting the partition $\mathcal{P}_B$ in $(i+1)$ parts. Any remaining hyperedge will contain terminals from B^c only and hence will intersect the partition $\mathcal{P}_B$ in only one part. As a result, we have $\sum_{e\in\mathcal{E}}[P_B(e)-1] = (t-1)\binom{|B|}{t} + \sum_{i=1}^{t-1}\binom{|B|}{i}\binom{m-|B|}{t-i} i = (t-1)\binom{|B|}{t} + |B|\sum_{i=1}^{t-1}\binom{|B|-1}{i-1}\binom{m-|B|}{t-i}$. Observe that $\sum_{i=1}^{t-1}\binom{|B|-1}{i-1}\binom{m-|B|}{t-i}$ is equal to $\binom{|B|-1}{t-1}$ subtracted from the coefficient of x^{t-1} in the expansion of $(1+x)^{|B|-1}(1+x)^{m-|B|} = (1+x)^{m-1}$. Therefore, $\sum_{i=1}^{t-1}\binom{|B|-1}{i-1}\binom{m-|B|}{t-i} = \binom{m-1}{t-1} - \binom{|B|-1}{t-1}$, and hence, for $|B| \geq t$, we have

$$\begin{aligned}&\sum_{e\in\mathcal{E}}[P_B(e)-1] \\ &\quad= |B|\binom{m-1}{t-1} + (t-1)\binom{|B|}{t} - |B|\binom{|B|-1}{t-1} \\ &\quad= |B|\binom{m-1}{t-1} - \binom{|B|}{t}.\end{aligned} \tag{49}$$

Next, we turn our attention to the case of $|B| < t$. In this case there are no hyperedges containing only terminals in B. For any i satisfying $1 \leq i \leq |B|$, there exist $\binom{|B|}{i}\binom{m-|B|}{t-i}$ hyperedges intersecting the partition in $(i+1)$ parts, as in the earlier case. However, all the remaining hyperedges are contained in B^c only, and hence play no part in the expression $\sum_{e\in\mathcal{E}}[P_B(e)-1]$. Thus, noting $|B| < t$, we have as in the

previous case,

$$\sum_{e\in\mathcal{E}}[P_B(e)-1] \;=\; |B|\sum_{i=1}^{|B|}\binom{|B|-1}{i-1}\binom{m-|B|}{t-i} \;=\; |B|\binom{m-1}{t-1}. \tag{50}$$

We will now apply Corollary 18. When $|B|\geq t$, using (49) we have

$$\begin{aligned}
&\frac{1}{|B|}\sum_{e\in\mathcal{E}}[P_B(e)-1]-\frac{(t-1)|\mathcal{E}|}{m-1}\\
&=\binom{m-1}{t-1}-\frac{1}{|B|}\binom{|B|}{t}-\frac{t-1}{m-1}\binom{m}{t} \qquad (51)\\
&=\frac{1}{t}\left[\frac{(m-1)!\;t}{(m-t)!\;(t-1)!}-\frac{m!}{(t-2)!\;(m-t)!\;(m-1)}\right.\\
&\qquad\left.-\binom{|B|-1}{t-1}\right]\\
&=\frac{1}{t}\left[\frac{(m-1)!}{(t-2)!\;(m-t)!}\left(\frac{t}{t-1}-\frac{m}{m-1}\right)-\binom{|B|-1}{t-1}\right]\\
&=\frac{1}{t}\left[\binom{m-2}{t-1}-\binom{|B|-1}{t-1}\right]\\
&>0, \qquad (52)
\end{aligned}$$

where (52) holds as $|B|\leq m-2$. For the case of $|B|<t$, we have

$$\begin{aligned}
\frac{1}{|B|}\sum_{e\in\mathcal{E}}[P_B(e)-1]-\frac{(t-1)|\mathcal{E}|}{m-1}&=\binom{m-1}{t-1}-\frac{t-1}{m-1}\binom{m}{t}\\
&=\frac{1}{t}\left[\binom{m-2}{t-1}\right] \qquad (53)\\
&>0,
\end{aligned}$$

where (53) follows from (51) and (52). Thus, using Corollary 18 we have the result. ∎

Next up is the proof of Corollary 20, which states that PIN models on k-regular, k-edge-connected graphs are strict Type $\mathcal{S}$.

Proof of Corollary 20: Consider a k-regular, k-edge-connected graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$. Using k-regularity, we have $|\mathcal{E}|=\frac{km}{2}$. As usual, we fix a $B\subsetneq\mathcal{M}$ satisfying $1\leq|B|\leq m-2$ and proceed to evaluate the expression $\sum_{e\in\mathcal{E}}[P_B(e)-1]$. Observe that for an ordinary graph, the sum $\sum_{e\in\mathcal{E}}[P_B(e)-1]=|\mathcal{E}_{\mathcal{P}_B}|$, where $\mathcal{E}_{\mathcal{P}_B}$ is the set of edges whose end-points lie in different cells of the partition $\mathcal{P}_B$. To proceed, we perform a graph contraction operation along the partition $\mathcal{P}_B$ on $\mathcal{G}$ to get a new graph $\mathcal{G}'=(\mathcal{V}',\mathcal{E}')$. More precisely, we take $\mathcal{V}'=B\bigcup\{B^c\}$ and $\mathcal{E}'=\left\{\{i,j\}\in\mathcal{E}:i,j\in B\right\}\bigcup\left\{\{B^c,i\}:\exists\{i,j\}\in\mathcal{E},i\in B,j\in B^c\right\}$, so that $|\mathcal{E}_{\mathcal{P}_B}|=|\mathcal{E}'|$. Now, the degree of every $v\in\mathcal{V}'$ satisfying $v\in B$ is k, whereas by the k-edge connectivity the degree of B^c in $\mathcal{G}'$ is at least k. Hence, we have $|\mathcal{E}_{\mathcal{P}_B}|=|\mathcal{E}'|\geq\frac{k(|B|+1)}{2}$. Therefore,

$$\begin{aligned}
\frac{1}{|B|}\sum_{e\in\mathcal{E}}[P_B(e)-1]-\frac{|\mathcal{E}|}{m-1}&=\frac{1}{|B|}|\mathcal{E}_{\mathcal{P}_B}|-\frac{km}{2(m-1)}\\
&\geq\frac{k}{2}\left[\frac{|B|+1}{|B|}-\frac{m}{m-1}\right]\\
&>0, \qquad (54)
\end{aligned}$$

where, (54) follows from the fact that $|B|\leq m-2$. Using Corollary 18 we have the result. ∎

Finally, we give the proof of Corollary 21, which states that a PIN model obtained from a Steiner triple system (STS) is strict Type $\mathcal{S}$. Recall that $\mathcal{H}_{\text{STS}}=(\mathcal{M},\text{STS}(\mathcal{M}))$ is a 3-uniform hypergraph obtained from an STS on $\mathcal{M}$.

Proof of Corollary 21: We will use Proposition 16 to get the result. First, we calculate $H(X_i)$ for any $i\in\mathcal{M}$. Observe that $H(X_i)$ counts the number of elements of STS($\mathcal{M}$) containing i. Now, fixing $i\in\mathcal{M}$, there are $m-1$ pairs of elements from $\mathcal{M}$ which contain i. Any set in STS($\mathcal{M}$) containing i contains two such pairs. Further, by the definition of STS, we know that any such pair is a subset of exactly one element of STS($\mathcal{M}$). Hence, we have $H(X_i)=\frac{m-1}{2}$. Next, we evaluate $H(X_{\mathcal{M}})=|\text{STS}(\mathcal{M})|$. Note that there are $\binom{m}{2}$ pairs of elements in $\mathcal{M}$, each pair being a subset of exactly one element of STS($\mathcal{M}$). Also, each element of STS($\mathcal{M}$) contains three such pairs. Therefore, we have $H(X_{\mathcal{M}})=|\text{STS}(\mathcal{M})|=\frac{m(m-1)}{6}$. Using these facts, we have $\Delta(\mathcal{S})=\frac{1}{m-1}\left[\frac{m(m-1)}{2}-\frac{m(m-1)}{6}\right]=\frac{m}{3}$.

Now, fix a $B\subsetneq\mathcal{M}$ with $1\leq|B|\leq m-2$ and evaluate $\Delta(\mathcal{P}_B)$. We consider two cases: $1\leq|B|\leq m-3$ and $|B|=m-2$. First, consider $1\leq|B|\leq m-3$. To proceed, we calculate a lower bound on $H(X_A)$ for any $A\subsetneq\mathcal{M}$. Observe that $H(X_A)$ counts the number of sets in STS($\mathcal{M}$) which contain at least one element from A. We will calculate an upper bound on the number of elements of STS($\mathcal{M}$) containing only elements of A^c, and subtract it from $|\text{STS}(\mathcal{M})|$ to get the required lower bound. The total number of pairs formed by the elements of A^c is $\binom{m-|A|}{2}$. Again, as each element of STS($\mathcal{M}$) contains 3 pairs, the required upper bound is $\lfloor\frac{(m-|A|)(m-|A|-1)}{6}\rfloor$. Thus, we have $H(X_A)\geq|\text{STS}(\mathcal{M})|-\frac{(m-|A|)(m-|A|-1)}{6}$. So, $H(X_{B^c})\geq|\text{STS}(\mathcal{M})|-\frac{|B|(|B|-1)}{6}$, and hence, $\Delta(\mathcal{P}_B)\geq\frac{1}{|B|}\left[\frac{|B|(m-1)}{2}-\frac{|B|(|B|-1)}{6}\right]=\frac{m-1}{2}-\frac{|B|-1}{6}$. Therefore,

$$\begin{aligned}
\Delta(\mathcal{P}_B)-\Delta(\mathcal{S})&\geq\frac{m-1}{2}-\frac{|B|-1}{6}-\frac{m}{3}\\
&=\frac{1}{6}[m-2-|B|]\\
&>0, \qquad (55)
\end{aligned}$$

where (55) follows from the fact that $|B|<m-2$.

To complete the proof, we show that $\Delta(\mathcal{P}_B)-\Delta(\mathcal{S})>0$ is satisfied when $|B|=m-2$. To this end, we fix a $B=\mathcal{M}\setminus\{i,j\}$, where $i,j\in\mathcal{M}$. We will exactly calculate $H(X_{B^c})$, which is the number of elements of STS($\mathcal{M}$) containing at least one of i and j. It has been shown earlier that i and j each occur in exactly $\frac{m-1}{2}$ elements, and they occur together exactly once. Therefore, we have $H(X_{B^c})=m-2$, and hence,

$\Delta(\mathcal{P}_B) = \frac{1}{m-2}\left[\frac{(m-2)(m-1)}{2} + (m-2) - \frac{m(m-1)}{6}\right]$. Thus,

$$\begin{aligned}\Delta(\mathcal{P}_B) - \Delta(\mathcal{S}) &= \frac{1}{m-2}\left[\frac{(m-2)(m-1)}{2} + (m-2) - \frac{m(m-1)}{6}\right] - \frac{m}{3} \\ &= \frac{m-3}{3(m-2)} \\ &> 0, \qquad (56)\end{aligned}$$

where (56) follows from the fact that $m > 3$. ■

ACKNOWLEDGEMENT

We would like to acknowledge the useful discussions we have had with Himanshu Tyagi, Chung Chan, Navid Nouri and Qiaoqiao Zhou that have deepened our understanding of multiterminal SK generation. We also thank the anonymous reviewers and the Associate Editor, Ashish Khisti, for their insightful comments that have helped improve the presentation of our work.

REFERENCES

[1] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, pp. 733–742, May 1993.

[2] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography, part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1121–1132, July 1993.

[3] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3047–3061, Dec. 2004.

[4] I. Csiszár and P. Narayan, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, June 2008.

[5] S. Nitinawarat, C. Ye, A. Barg, P. Narayan and A. Reznik, "Secret key generation for a pairwise independent network model," *IEEE Trans. Inf. Theory*, vol. 56, pp. 6482–6489, Dec. 2010.

[6] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy and Steiner tree packing," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6490–6500, Dec. 2010.

[7] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals–Part I," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973–3996, Aug. 2010.

[8] H. Tyagi, "Common information and secret key capacity," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5627–5640, Sep. 2013.

[9] A. C. Yao, "Some complexity questions related to distributed computing," in *Proc. 11th Annu. ACM Symp. Theory of Computing (STOC)*, 1979.

[10] A. D. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 163–179, Mar. 1975.

[11] M. Mukherjee, N. Kashyap and Y. Sankarasubramaniam, "Achieving SK capacity in the source model: When must all terminals talk?," in *Proc. 2014 IEEE Int. Symp. Inf. Theory (ISIT 2014)*, Honolulu, Hawai'i, USA, June 29 – July 4, 2014, pp. 1156–1160.

[12] C. Chan, A. Al-Bashabsheh, J. Ebrahimi, T. Kaced and T. Liu, "Multivariate mutual information inspired by secret key agreement," in *Proc. of IEEE*, vol. 103, no. 10, pp. 1883-1913, Oct. 2015.

[13] H. Zhang, Y. Liang and L. Lai, "Secret key capacity: Talk or keep silent?," in *Proc. 2015 IEEE Int. Symp. Inf. Theory (ISIT 2015)*, Hong Kong, China, June 14–19, 2015, pp. 291–295.

[14] T. A. Courtade and R. D. Wesel, "Coded cooperative data exchange in multihop networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1136–1158, Feb. 2014.

[15] T. A. Courtade and T. R. Halford, "Coded cooperative data exchange for a secret key," in *Proc. 2014 IEEE Int. Symp. Inf. Theory (ISIT 2014)*, Honolulu, Hawai'i, USA, June 29 – July 4, 2014, pp. 776–780.

[16] T. A. Courtade and T. R. Halford, "Coded cooperative data exchange for a secret key," Arxiv:1407.0333v1.

[17] S. El Rouayheb, A. Sprintson, and P. Sadeghi, "On coding for cooperative data exchange," in *Proc. 2010 IEEE Inf. Theory Workshop (ITW 2010)*, Cairo, Egypt, 6–8 Jan. 2010, pp. 1–5.

[18] J. Liu, P. Cuff and S. Verdu, "Secret key generation with one communicator and a strong converse via hypercontractivity", in *Proc. 2015 IEEE Int. Symp. Inf. Theory (ISIT 2015)*, Hong Kong, China, June 14–19, 2015, pp. 710–714.

[19] L. Zhao, "Common Randomness, Efficiency, and Actions", *PhD thesis*, Stanford University, 2011.

[20] T.A. Courtade, "Outer bounds for multiterminal source coding via a strong data processing inequality", *Proc. 2013 IEEE Int. Symp. Inf. Theory (ISIT 2013)*, Istanbul, Turkey, July 7–12, 2013, pp. 559–563.

[21] V. Anantharam, A. Gohari, S. Kamath, and C. Nair, "On hypercontractivity and a data processing inequality," *Proc. 2014 IEEE Int. Symp. Inf. Theory (ISIT 2014)*, Honolulu, Hawaii, USA, June 29 – July 4, 2014, pp. 3022–3026.

[22] M. Braverman and A. Rao, "Information equals amortized communication," *IEEE Trans. Inf. Theory*, vol. 60, pp. 6058–6069, Oct. 2014.

[23] M. Braverman and J. Schneider, "Information complexity is computable," *Electronic Colloquium on Computational Complexity (ECCC)*, Report No. 23, 2015.

[24] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2699–2713, June 2010.

[25] C. Chan and L. Zheng, "Mutual dependence for secret key agreement," in *Proc. 44th Annual Conference on Information Sciences and Systems (CISS)*, 2010.

[26] M.R. Bloch and J.N. Laneman, "Strong secrecy from channel resolvability," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.

[27] G. Xu, W. Liu and B. Chen, "Wyner's common information: Generalizations and a new lossy source coding interpretation," Arxiv:1301.2237v1.

[28] R. Tandon, L. Sankar and H.V. Poor, "Multi-user privacy: The Gray-Wyner system and generalized common information," in *Proc. 2011 IEEE Int. Symp. Inf. Theory (ISIT 2011)*, St. Petersburg, Russia, July 31 – Aug. 5, 2011, pp. 563–567.

[29] A. El Gamal and Y. H. Kim, *Network Information Theory*, Cambridge University Press, 2011.

[30] C. Chan, A. Al-Bashabsheh, J. Ebrahimi, T. Kaced, S. Kadhe, T. Liu, A. Sprintson, M. Yan and Q. Zhou, "Successive Omniscience," in *Proc. 2015 Int. Symp. Network Coding (NetCod 2015)*, Sydney, Australia, June 22–24, 2015, pp. 21-25.

[31] A. M. Odlyzko, "Asymptotic enumeration methods," in *Handbook of Combinatorics*, R.L. Graham et al., eds., 1995, pp. 1063–1229.

[32] H. Tyagi, N. Kashyap, Y. Sankarasubramaniam and K. Viswanathan, "Fault tolerant secret key generation," in *Proc. 2012 IEEE Int. Symp. Inf. Theory (ISIT 2012)*, Cambridge, Massachusetts, USA, July 1–6, 2012, pp. 1787–1791.

[33] F. Harary, "Maximum connectivity of a graph," in *Proc. Nat. Acad. Sci.*, vol. 48, pp. 1142–1145, 1962.

[34] N. Kashyap, M. Mukherjee and Y. Sankarasubramaniam, "On the secret key capacity of the Harary graph PIN model," in *Proc. 2013 Nat. Conf. Commun. (NCC 2013)*, Delhi, India, Feb. 15–17, 2013, pp. 1–5.

[35] C. J. Colbourn and A. Rosa, *Triple Systems*, Oxford Mathematical Monographs, 1999.

Manuj Mukherjee (S'14) received the B.E. degree in Electronics and Telecommunication Engineering from the Jadavpur University, Kolkata, in 2011. Since 2011, he has been a Ph.D. student in the department of Electrical Communication Engineering, Indian Institute of Science, Bangalore. His research interests are information theoretic security, secret key generation and multiterminal information theory.

Navin Kashyap (S'97-M'02-SM'07) received the B.Tech. degree in Electrical Engineering from the Indian Institute of Technology, Bombay, in 1995, the M.S. degree in Electrical Engineering from the University of Missouri-Rolla in 1997, and the M.S. degree in Mathematics and the Ph.D. degree in Electrical Engineering from the University of Michigan, Ann Arbor, in 2001. From November 2001 to November 2003, he was a postdoctoral research associate at the University of California, San Diego. From 2004 to 2010, he was at the Department of Mathematics and Statistics at Queen's University, Kingston, Ontario, first as an Assistant Professor, then as an Associate Professor. In January 2011, he joined the Department of Electrical Communication Engineering at the Indian Institute of Science as an Associate Professor. His research interests lie primarily in the application of combinatorial and probabilistic methods in information and coding theory. Prof. Kashyap served on the editorial board of the IEEE Transactions on Information Theory during the period 2009–2014.

Yogesh Sankarasubramaniam received the B.Tech degree from the Indian Institute of Technology, Madras (2001), and the M.S. (2003) and Ph.D. (2006) degrees from Georgia Institute of Technology, Atlanta, in Electrical Engineering. He has held positions as a Research Scientist at HP Labs India and IBM Research India. His interests are in information theory and applied mathematics.